

REFERENCE: B05

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 August 2006 (10.08.2006)

PCT

(10) International Publication Number
WO 2006/084272 A2

- (51) International Patent Classification:
C12Q 1/68 (2006.01)
- (21) International Application Number:
PCT/US2006/004280
- (22) International Filing Date: 6 February 2006 (06.02.2006)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/650,365 4 February 2005 (04.02.2005) US

(71) Applicants (for all designated States except US):
ROSETTA INPHARMATICS LLC [US/US], 401 Terry
Avenue North, Seattle, Washington 98108 (US); MERCK
& CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway,
New Jersey 07065 (US).

- (72) Inventors; and
(75) Inventors/Applicants (for US only): DAI, Hongyue
[CN/US]; 16814 118th Avenue N.e., Bothell, Washington

98011 (US); FRIEND, Stephen, H [US/US]; 101 W.
Mermaid Lane, Philadelphia, Pennsylvania 19118 (US);
DEUTSCH, Paul [US/US]; 50 Maybury Hill Road,
Princeton, New Jersey 08540 (US).

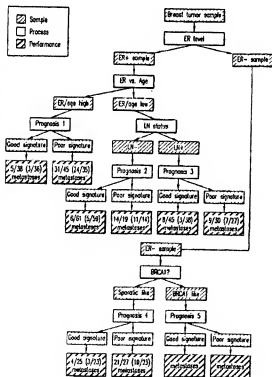
(74) Agents: ANTILER, Adriane, M. et al.; JONES DAY, 222
East 41st Street, New York, New York 10017-6702 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DR, DZ, EC, EG, ES, FI,
GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE,
KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV,
LY, MA, MD, MG, MK, MN, MW, MX, MY, NA, NG, NI,
NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG,
SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM,
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

[Continued on next page]

(54) Title: METHODS OF PREDICTING CHEMOTHERAPY RESPONSIVENESS IN BREAST CANCER PATIENTS



(57) Abstract: The invention provides a method of predicting the responsiveness of a breast cancer patient to chemotherapy based on the patient's cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from the patient, the patient's estrogen receptor level, and the patient's estrogen receptor level relative to the patient's age. The invention also provides a method for selecting patients for enrollment in a clinical trial of a drug for treating breast cancer based on these factors.

WO 2006/084272 A2



FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *without international search report and to be republished upon receipt of that report*

METHODS OF PREDICTING CHEMOTHERAPY RESPONSIVENESS IN BREAST CANCER PATIENTS

This application claims benefit of U.S. Provisional Patent Application No.
60/650,365, filed on February 4, 2005, which is incorporated by reference herein in its
entirety.

1. FIELD OF THE INVENTION

The present invention relates to a method of predicting chemotherapy responsiveness
in breast cancer patients. The invention also relates to a method of selecting patients for
enrollment in clinical trials of breast cancer drugs.

2. BACKGROUND OF THE INVENTION

The increased number of cancer cases reported in the United States, and, indeed,
around the world, is a major concern. Currently there are only a handful of treatments
available for specific types of cancer, and these provide no guarantee of success. In order to
be most effective, these treatments require not only an early detection of the malignancy, but
a reliable assessment of the severity of the malignancy.

The incidence of breast cancer, a leading cause of death in women, has been gradually
increasing in the United States over the last thirty years. Its cumulative risk is relatively high;
1 in 8 women are expected to develop some type of breast cancer by age 85 in the United
States. In fact, breast cancer is the most common cancer in women and the second most
common cause of cancer death in the United States. In 1997, it was estimated that 181,000
new cases were reported in the U.S., and that 44,000 people would die of breast cancer
(Parker *et al.*, *CA Cancer J. Clin.* 47:5-27 (1997); Chu *et al.*, *J. Nat. Cancer Inst.* 88:1571-
1579 (1996)). While mechanism of tumorigenesis for most breast carcinomas is largely
unknown, there are genetic factors that can predispose some women to developing breast
cancer (Miki *et al.*, *Science*, 266:66-71(1994)).

Sporadic tumors, those not currently associated with a known germline mutation,
constitute the majority of breast cancers. It is also likely that other, non-genetic factors also
have a significant effect on the etiology of the disease. Regardless of the cancer's origin,
breast cancer morbidity and mortality increases significantly if it is not detected early in its

progression. Thus, considerable effort has focused on the early detection of cellular transformation and tumor formation in breast tissue.

A marker-based approach to tumor identification and characterization promises improved diagnostic and prognostic reliability. Typically, the diagnosis of breast cancer requires histopathological proof of the presence of the tumor. In addition to diagnosis, histopathological examinations also provide information about prognosis and selection of treatment regimens. Prognosis may also be established based upon clinical parameters such as tumor size, tumor grade, the age of the patient, and lymph node metastasis.

Diagnosis and/or prognosis may be determined to varying degrees of effectiveness by direct examination of the outside of the breast, or through mammography or other X-ray imaging methods (Jatoi, *Am. J. Surg.* 177:518-524 (1999)). The latter approach is not without considerable cost, however. Every time a mammogram is taken, the patient incurs a small risk of having a breast tumor induced by the ionizing properties of the radiation used during the test. In addition, the process is expensive and the subjective interpretations of a technician can lead to imprecision. For example, one study showed major clinical disagreements for about one-third of a set of mammograms that were interpreted individually by a surveyed group of radiologists. Moreover, many women find that undergoing a mammogram is a painful experience. Accordingly, the National Cancer Institute has not recommended mammograms for women under fifty years of age, since this group is not as likely to develop breast cancers as are older women. It is compelling to note, however, that while only about 22% of breast cancers occur in women under fifty, data suggests that breast cancer is more aggressive in pre-menopausal women.

In clinical practice, accurate diagnosis of various subtypes of breast cancer is important because treatment options, prognosis, and the likelihood of therapeutic response all vary broadly depending on the diagnosis. Accurate prognosis, or determination of distant metastasis-free survival could allow the oncologist to tailor the administration of adjuvant chemotherapy, with women having poorer prognoses being given the most aggressive treatment. Furthermore, accurate prediction of poor prognosis would greatly impact clinical trials for new breast cancer therapies, because potential study patients could then be stratified according to prognosis. Trials could then be limited to patients having poor prognosis, in turn making it easier to discern if an experimental therapy is efficacious.

To date, no set of satisfactory predictors for prognosis based on the clinical information alone has been identified. Many have observed that the ER status has a dominant signature in the breast tumor gene expression profiling. See West *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 98:11462 (2001); van 't Veer *et al.*, *Nature* 415:530 (2002); Sorlie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:8418 (2003); Perou *et al.*, *Nature* 406:747 (2000); Gruvberger *et al.*, *Cancer Res.* 61:5979 (2001); Sotiriou *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:10393 (2003). It is generally accepted that there is some relationship between patient survival and ER status. van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002); Surowiak *et al.*, *Folia Histochem. Cytobiol.* 39:143 (2001); Pichon *et al.*, *Br. J. Cancer* 73:1545 (1996); Collett *et al.*, *J. Clin. Pathol.* 49:920 (1996). *BRCA1* mutations are related to the familial cancer susceptibility. Biesecker *et al.*, *JAMA* 269:1970 (1993); Easton *et al.*, *Cancer Surv.* 18:95 (1993). Age is also considered to be a prognosis factor since young cancer patients tend to have poor tumors. Maggard *et al.*, *J. Surg. Res.* 113:109 (2003). Lymph node status is a factor in deciding the treatment. Eifel *et al.*, *J. Natl. Cancer Inst.* 93:979 (2001).

The discovery and characterization of *BRCA1* and *BRCA2* has recently expanded our knowledge of genetic factors which can contribute to familial breast cancer. Germ-line mutations within these two loci are associated with a 50 to 85% lifetime risk of breast and/or ovarian cancer (Casey, *Curr. Opin. Oncol.* 9:88-93 (1997); Marcus *et al.*, *Cancer* 77:697-709 (1996)). Only about 5% to 10% of breast cancers, however, are associated with breast cancer susceptibility genes, *BRCA1* and *BRCA2*. The cumulative lifetime risk of breast cancer for women who carry the mutant *BRCA1* is predicted to be approximately 92%, while the cumulative lifetime risk for the non-carrier majority is estimated to be approximately 10%. *BRCA1* is a tumor suppressor gene that is involved in DNA repair and cell cycle control, which are both important for the maintenance of genomic stability. More than 90% of all mutations reported so far result in a premature truncation of the protein product with abnormal or abolished function. The histology of breast cancer in *BRCA1* mutation carriers differs from that in sporadic cases, but mutation analysis is the only way to find the carrier. Like *BRCA1*, *BRCA2* is involved in the development of breast cancer, and like *BRCA1* plays a role in DNA repair. However, unlike *BRCA1*, it is not involved in ovarian cancer.

Other genes have been linked to breast cancer, for example c-erb-2 (*HER2*) and p53 (Beenken *et al.*, *Ann. Surg.* 233(5):630-638 (2001)). Overexpression of c-erb-2 (*HER2*) and p53 have been correlated with poor prognosis (Rudolph *et al.*, *Hum. Pathol.* 32(3):311-319

(2001), as has been aberrant expression products of *mdm2* (Lukas *et al.*, Cancer Res. 61(7):3212-3219 (2001) and cyclin1 and p27 (Porter & Roberts, International Publication WO98/33450, published August 6, 1998).

The detection of *BRCA1* or *BRCA2* mutations represents a step towards the design of therapies to better control and prevent the appearance of these tumors. Recently, many studies have used gene expression profiling to analyze various cancers, and those studies have provided new diagnosis and prognosis information in the molecular level. See Zajchowski *et al.*, "Identification of Gene Expression Profiled that Predict the Aggressive Behavior of Breast Cancer Cells," *Cancer Res.* 61:5168 (2001); West *et al.*, "Predicting the Clinical Status of Human Breast Cancer by Using Gene Expression Profiles," *Proc. Natl. Acad. Sci. U.S.A.* 98:11462 (2001); van 't Veer *et al.*, "Gene Expression Profiling Predicts the Outcome of Breast Cancer," *Nature* 415:530 (2002); Roberts *et al.*, "Diagnosis and Prognosis of Breast Cancer Patients," WO 02/103320; Sorlie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:8418 (2003); Perou *et al.*, *Nature* 406:747 (2000); Khan *et al.*, *Cancer Res* 58, 5009 (1998); Golub *et al.*, *Science* 286, 531 (1999); DeRisi *et al.*, *Nat. Genet.* 14:457 (1996); Alizadeh *et al.*, *Nature* 403, 503 (2000). Methods for the identification of informative genesets for various cancers have also been described. See Roberts *et al.*, "Diagnosis and Prognosis of Breast Cancer Patients," WO 02/103320; Golub *et al.*, United States Patent No. 6,647,341.

Genesets have been identified that are informative for differentiating individuals having, or suspected of having, breast cancer based on estrogen receptor (ER) status, or *BRCA1* mutation vs. sporadic (*i.e.*, other than *BRCA1*-type) mutation status. See Roberts *et al.*, WO 02/103320; van't Veer *et al.*, *Nature* 415:530 (2001). Genesets have also been identified that enable the classification of sporadic tumor-type individuals as those who will likely have no metastases within five years of initial diagnosis (*i.e.*, individuals with a good prognosis) or those who will likely have a metastasis within five years of initial diagnosis (*i.e.*, those having a poor prognosis). Roberts, *supra*; van't Veer, *supra*.

Roberts *et al.* WO 02/103320 and van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002) describe a 70-gene set, useful for the prognosis of breast cancer, which outperformed clinical measures of prognosis, and which showed good potential in selecting good outcome patients, thereby avoiding over-treatment. The expression of genes with most predictive

value, however, was not homogeneous among poor patients, suggesting the need for improvement.

Although the patterns of gene expression as described in Roberts *et al.* were correlated with existing clinical indicators such as estrogen receptor and *BRCA1* status, clinical measures were not incorporated. Furthermore, although the poor-outcome group in particular showed heterogeneity in expression pattern, the best classifier decision rule found during these studies was a fairly simple one based on the similarity of a patient profile to the average profile of a good-outcome training group.

It is evident that breast cancer is the result of more than one type of molecular event. Likewise, a variety of other conditions, such as other cancers; non-cancer diseases such as diabetes, autoimmune or neurodegenerative disorders, obesity; etc., are also the result of more than one molecular event. Moreover, an individual's response to exposure to particular environmental conditions, for example, exposure to natural or man-made agents, such as toxins, pollutants, drugs, food additives, etc., likely result from more than one molecular event. Thus, there exists a need for improved prognostic methods so that appropriate courses of prophylaxis and/or therapy may be provided. Genesets having improved prognostic power can be identified by first identifying discrete subsets of individuals based on genotypic or phenotypic characteristics relevant to the disease or condition, and then identifying genesets informative for prognosis within those subsets of patients. Individuals having the condition, or who are suspected of having the condition, such as breast cancer, would then be provided therapies appropriate to the molecular mechanisms underlying the condition. The present invention provides such methods for breast cancer, and for other cancers, diseases or conditions.

3. SUMMARY OF THE INVENTION

The present invention provides a method for predicting responsiveness of a breast cancer patient to chemotherapy based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method

comprising predicting said patient to exhibit (a) enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is low; or (b) reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if (i) said ER level is ER⁻, or (ii) said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high; wherein said cellular constituent profile is classified as a poor prognosis profile if said cellular constituent profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of good outcome patients and said poor prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-occurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a patient who has occurrence of metastases within a second period of time after initial diagnosis.

In one embodiment, said cellular constituent profile is determined as a poor prognosis profile, and said ER level is determined as ER⁺, and said ER/AGE is determined as low, and said patient is predicted to have enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients. In another embodiment, said cellular constituent profile is determined as a poor prognosis profile, said ER level is determined as ER⁺, and said ER/AGE is determined as high, and said patient is predicted to have reduced response to chemotherapy as compared to patients in the general population of breast cancer patients. In still another embodiment, said ER level is determined as ER⁻, and said patient is predicted to have reduced response to chemotherapy as compared to patients in the general population of breast cancer patients.

The invention also provides a method for selecting a patient for enrollment in a clinical trial of a drug for treating breast cancer based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said

ER/AGR being a metric of said ER level relative to the age of said patient, said method comprising selecting a patient for inclusion in said clinical trial if (a) said ER level is ER⁻; or (b) said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high; wherein said cellular constituent profile is classified as a poor prognosis
5 profile if said cellular constituent profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of good outcome patients and said poor prognosis template comprising measurements of said plurality of cellular constituents representative of levels of
10 said cellular constituents in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-occurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a breast cancer patient who has occurrence of metastases within a second period of time after initial diagnosis.

In one embodiment, said ER level is determined as ER⁻, and said patient is selected.

15 In another embodiment, said cellular constituent profile is determined as a poor prognosis profile, said ER level is determined as ER⁺, and said ER/AGE is determined as high, and said patient is selected.

The invention also provides a method for identifying a breast cancer patient as a good or poor candidate for chemotherapy based on one or more of the following (i) a cellular
20 constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method
25 comprising (a) determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, said breast cancer patient is identified as a good candidate for chemotherapy; or (b) determining whether said ER level is ER⁻, whereby if said ER level is ER⁻, said the breast cancer patient is
30 identified as a poor candidate for chemotherapy; or (c) determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺,

and said ER/AGE is high, said breast cancer patient is identified as a poor candidate for chemotherapy.

5 In one embodiment, said cellular constituent profile is determined as a poor prognosis profile, said ER level is determined as ER⁺, and said ER/AGE is determined as low, and said breast cancer patient is identified as a good candidate for chemotherapy. In another embodiment, said ER level is determined as ER⁻, and said the breast cancer patient is identified as a poor candidate for chemotherapy. In still another embodiment, said cellular constituent profile is determined as a poor prognosis profile, said ER level is determined as ER⁺, and said ER/AGE is determined as high, and said breast cancer patient is identified as a
10 poor candidate for chemotherapy.

In embodiments of the methods of the invention, said first period is 10 years and said second period is 10 years. In another embodiment, said patient is under the age of 55, and wherein each said good outcome patient and each said poor outcome patient are under the age of 55 at time of diagnosis of breast cancer.

15 In one embodiment, any one of the methods of the invention further comprises determining said cellular constituent profile, said ER level, and/or, said ER/AGE.

In another embodiment, in any one of the methods of the invention said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient, wherein said good prognosis template comprises measurements of said
20 plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

In one embodiment, said cellular constituent profile is a differential expression profile
25 comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample. In one embodiment, said differential measurements is selected from the group consisting of xdev, log(ratio), error-weighted log(ratio), and mean subtracted log(intensity).

In another embodiment, said cellular constituent profile comprises measurements of a
30 plurality of protein species in a sample derived from said patient, wherein said good

prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome patients.

- 5 In one embodiment, measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

- In one embodiment, similarity of said cellular constituent profile to said good prognosis template is represented by a correlation coefficient between said cellular
10 constituent profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold, e.g., 0.5, indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

- In another embodiment, similarity of said cellular constituent profile to said good prognosis template is represented by a distance between said cellular constituent profile and
15 said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance equal to or greater than said given value indicates a low similarity.

- In one embodiment, said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene
20 in said control sample, and wherein said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and wherein said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65. In one embodiment, said gene encoding said estrogen receptor is the estrogen receptor α gene.

- In one embodiment, said ER/AGE is classified as high if said ER level as measured
25 by $\log_{10}(\text{ratio})$ is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

 In one embodiment, said estrogen receptor level is measured by an oligonucleotide probe that detects a transcript corresponding to the gene having accession number

NM_000125, wherein said control sample is a pool of breast cancer cells of different patients, and wherein $c = 0.1$ and $d = 42.5$.

- In one embodiment, said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients. In another embodiment,
- 5 said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript of said gene encoding said estrogen receptor.

In one embodiment, said plurality of transcripts consists of transcripts corresponding to at least a portion of the set of genes listed in Table 8. In another embodiment, said plurality of transcripts consists of all transcripts corresponding to genes listed in Table 8.

- 10 In one embodiment, said chemotherapy is carried out using the CMF combination consisting of cyclophosphamide, methotrexate, and 5-fluorouracil.

The invention also provides a computer system comprising a processor, and a memory coupled to said processor and encoding one or more programs that cause the processor to carry out any one of the methods of the invention.

- 15 The invention also provides computer program product for use in conjunction with a computer having a processor and a memory connected to the processor, said computer program product comprising a computer readable storage medium having a computer program mechanism encoded thereon, which computer program mechanism may be loaded into the memory of said computer and cause said computer to carry out any one of the
- 20 methods of the invention.

- The invention also provides a method for predicting responsiveness of a breast cancer patient to chemotherapy based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor
- 25 transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising predicting said patient to exhibit (a) enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular
- 30 constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is

low; or (b) reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if (i) said ER level is ER⁻, or (ii) said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high, wherein said cellular constituent profile is classified as a good prognosis profile if said
5 cellular constituent profile predicts non-occurrence of metastases in said breast cancer patient within a predetermined period of time after initial diagnosis, and wherein said cellular constituent profile is classified as a poor prognosis profile if said cellular constituent profile predicts occurrence of metastasis within said predetermined period of time. In one embodiment, the predetermined period of time is 5 or 10 years.

10 In one embodiment, said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient. In another embodiment, said cellular constituent profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample. In one embodiment, said
15 control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts.

In a preferred embodiment, said plurality of transcripts consists of transcripts corresponding to at least a portion, e.g., at least 5, 10, 40, 50 or all, of the genes listed in Table 8.

20 In one embodiment, said cellular constituent profile is classified as a good prognosis profile or a poor prognosis profile using an artificial neural network, which receives said cellular constituent profile as an input and generates an output comprising data indicating whether said cellular constituent profile is a good prognosis profile or a poor prognosis profile.

25 In another embodiment, said cellular constituent profile is classified as a good prognosis profile or a poor prognosis profile using a supporting vector machine, which receives said cellular constituent profile as an input and generates an output comprising data indicating whether said cellular constituent profile is a good prognosis profile or a poor prognosis profile.

30 4. BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 depicts the decision tree that resulted in the five patient subsets used to identify informative prognosis-related genes.

FIG. 2 shows the relationship between ER level and age. (A) Scatter plot of ER vs. age for ER+ patients. Black dots indicate metastases free samples, and gray dots indicate metastases samples. It appears that patients of ER+ group can be subdivided into "ER+, ER/AGE high" group (above the black line) and "ER+, ER/AGE low" (below the black line) group. The black line is approximated by $ER = 0.1 * (AGE - 42.5)$, and the dashed line by $ER = 0.1 * (age - 50)$. Within each population, the ER level also increases with age. (B) Age distribution of all patients in ER+ samples. A bimodal distribution is observed. (C) ER-modulated age ($age - 10^*$) distribution of all patients in ER+ samples. A bimodal distribution is observed. (D) Age distribution of samples with metastasis. (E) ER-modulated age distribution of samples with metastasis. The three peaks appearing in this distribution suggest a polymorphism.

FIG. 3. Performance of classifier for the "ER+/sporadic" group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 4. Performance of classifier for the "ER+, ER/AGE high" group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, and squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 5. Performance of classifier for the "ER+, ER/AGE low/LN⁻" group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot

between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, and squares indicates samples with metastases. Dashed line indicates the threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 6. Performance of classifier for the "ER+, ER/AGE low/LN+" group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 7. Performance of classifier for the "ER-, BRCA1" group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 8. Heatmaps of genes representing key biological functions in subgroups of patients: A: Cell cycle genes are predictive of outcome in patients with ER/age high. B: Cell cycle genes are not predictive of outcome in "ER- and sporadic" patients C: Glycolysis genes are predictive of outcome in patients with ER/age low and LN-. D: Glycolysis genes are not predictive of outcome in "ER- & BRCA1" patients.

FIG. 9 shows metastasis free probability for the entire set of breast cancer patients younger than 55 years of age, with or without chemotherapy.

FIG. 10 shows the effect of chemotherapy in (a) patients with "good signature" as predicted by the "70 gene" classifier, and (b) patients with "poor signature".

FIG. 11 shows the effect of chemotherapy in (a) ER- patients, and (b) ER+ patients with "poor signature".

5 FIG. 12 ER+ patients are divided into "ER/age high" (above the line) and "ER/age low" (below the line) group according to their ER expression level relative to patient age at diagnosis. The ER expression level is measured by gene expression array. The control channel is the pool of all breast cancer samples. The line is represented as $y = 0.1 * (x - 42.5)$.

10 FIG. 13 shows the effect of chemotherapy in "poor signature" patients in (a) "ER/age high" group, and (b) "ER/age low" group.

5. DETAILED DESCRIPTION OF THE INVENTION

The invention provides a method of predicting the responsiveness of a breast cancer patient to chemotherapy based on the patient's cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from the patient, the patient's estrogen receptor level (hereinafter designated as "ER level"), and the patient's estrogen receptor level relative to the patient's age (hereinafter designated as "ER/AGE"). In the method of the invention, a patient is classified into an appropriate chemotherapy responsiveness group as compared to patients in the general population of breast cancer patients. The inventors have discovered that, among patients whose cellular constituent profile indicates poor prognosis, a patient's responsiveness to chemotherapy depends not only on the patient's ER level, but also on the change of the ER level with age. The inventors have found that patients who show high ER level at an earlier age (thus a high ER/AGE) show little response to chemotherapy, whereas patients who show high ER level at later age (thus a low ER/AGE) show increased response to chemotherapy. In particular, the inventors have found that a patient exhibits enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if the patient's cellular constituent profile indicates poor prognosis, and the patient's ER level is ER⁺, and the patient's estrogen receptor level relative to the patient's age ER/AGE is low, whereas a patient exhibits reduced response to chemotherapy as compared to patients in the general population if (i) the patient's ER level is ER⁺, or (ii) the patient's cellular constituent profile indicates poor prognosis, and the patient's ER level is ER⁺, and the patient's estrogen receptor level relative to the patient's

age ER/AGE is high. Herein, ER⁺ designates a high ER level and ER⁻ designates a low ER level. In addition, the inventors have also found that if a patient's cellular constituent profile indicates a good prognosis, then the patient does not need chemotherapy. The method of the invention is particularly useful for predicting the responsiveness of a breast cancer patient
5 under the age of 55.

The invention also provides a method for selecting patients for enrollment in a clinical trial of a drug for treating breast cancer based on each patient's cellular constituent profile, the patient's estrogen receptor level, and the patient's estrogen receptor level relative to the patient's age. In the method of the invention, patients having reduced responsiveness to
10 chemotherapy as compared to patients in the general population of breast cancer patients are preferably selected for inclusion in a clinical trial. In one embodiment, a patient who has either ER⁻ or a combination of (a) poor prognosis, (b) ER⁺, and (c) ER/AGE high are selected for inclusion in the trial. Such patients are predicted not to exhibit significant response to chemotherapy, and thus are good candidates in determining efficacy of a new breast cancer
15 drug.

Patient groups can be classified according to at least one of age, lymph node (LN) status, estrogen receptor (ER) level, and *BRCA1* mutation status into discrete patient subsets. These clinical factors have been implicated in tumor etiology as well as differences in disease outcome. These characteristics are not limiting; other genotypic or phenotypic characteristics
20 of breast cancer, for example, tumor grade, tumor size, tumor cell type, etc., may be used, alone or in combination with those listed herein, in order to classify individuals. The differences in gene expression or in tumor fate related to these parameters likely represent differences in tumor origin and tumor genesis, and are therefore good candidates for tumor stratification. Genesets informative for prognosis within each subset are then identified.
25 New breast cancer patients are then classified using the same criteria, and a prognosis is made based on the geneset specific for the patient subset into which the patient falls. In the process of constructing a prognosis classifier within each patient subset, particular attention is paid to the homogeneous patterns related to the tumor outcome. Emergence of such homogeneous prognosis patterns may indicate the most common mechanism to metastasis within a subset.
30 At the same time, successful identification of such patterns also justifies the parameters being used for the tumor stratification. To differentiate this approach from an mRNA-alone

approach, we refer to the current approach of integrating clinical data with the gene expression data as a "comprehensive prognosis".

The measurements in the cellular constituent profiles can be any suitable measured values of the cellular constituents, e.g., measurements of the expression levels of genes. The measurement of the expression level of a gene can be direct or indirect, e.g., directly of abundance levels of RNAs or proteins or indirectly, by measuring abundance levels of cDNAs, amplified RNAs or DNAs, proteins, or activity levels of RNAs or proteins, or other molecules (e.g., a metabolite) that are indicative of the foregoing. In one embodiment, the profile comprises measurements of abundances of the transcripts of the marker genes. The measurements of transcripts can be direct (of the transcripts themselves) or indirect (of, e.g., their cDNAs). The measurement of abundance can be a measurement of the absolute abundance of a gene product. The measurement of abundance can also be a value representative of the absolute abundance, e.g., a normalized abundance value (e.g., an abundance normalized against the abundance of a reference gene product) or an averaged abundance value (e.g., average of abundances obtained at different time points or from different tumor cell samples from the patients, or average of abundances obtained using different probes, etc.), or a combination of both. As an example, the measurement of abundance of a gene transcript can be a value obtained using an Affymetrix® GeneChip® to measure hybridization to the transcript.

In another embodiment, the expression profile is a differential expression profile comprising differential measurements of a plurality of transcripts in a sample derived from the patient versus measurements of the plurality of transcripts in a reference sample, e.g., a cell sample of normal cells. Each differential measurement in the profile can be but is not limited to an arithmetic difference, a ratio, or a log(ratio). As an example, the measurement of abundance of a gene transcript can be a value for the transcript obtained using a cDNA or ink-jet polynucleotide array in a two-color measurement.

As used herein, "*BRCA1* tumor" or "*BRCA1* type" means a tumor having cells containing a mutation of the *BRCA1* locus.

A "patient subset" is a group of individuals, all of whom have a particular condition, or are subject to a particular condition, which is distinguished from other individuals having that condition by one or more phenotypic, genotypic or clinical characteristics of the

condition, or of a response to the condition. For example, where the condition is breast cancer, individuals may belong to an "ER⁺" or an "ER⁻" patient subset, or may belong to a particular age group patient subset.

A gene and/or marker is "informative" for a condition, phenotype, genotype or clinical characteristic if the expression of the gene or marker is correlated or anticorrelated with the condition, phenotype, genotype or clinical characteristic to a greater degree than would be expected by chance.

The patient's estrogen receptor level relative to the patient's age can be measured using an appropriate metric designated as ER/AGE. An individual of a given age can be classified as "ER/AGE high" if the individual's ER level is higher than a threshold value for the given age. The threshold can be age-dependent, i.e., a different threshold for each different age. In one embodiment, the age-dependent threshold value is calculated as $c \cdot (AGE - d)$, where c is a coefficient, AGE is the age of the patient, and d is an age threshold. The parameters c and d depend on the particular measures and/or units of the ER level and AGE. They can be determined by fitting patients' ER level-age distribution to a bimodal distribution of two subgroups each having a different ER level-age dependence. In a specific embodiment, $c = 0.1$ and $d = 42.5$ is used for ER levels represented by a log(ratio) of ER expression level. Thus, for example, the threshold for a 45-year old individual in this embodiment is $0.1 (45 - 42.5)$, or 0.25, and if the log(ratio) of ER expression level of the individual is equal to or greater than 0.25, the individual is classified as "ER/AGE high"; otherwise, the individual is classified as "ER/AGE low."

5.1 IDENTIFICATION OF DIAGNOSTIC AND PROGNOSTIC MARKER SETS

The sets of genes and/or markers that can be used in conjunction with the present invention for diagnosis and/or prognosis of breast cancer can be identified using methods described in this section. In preferred embodiments, the method involves first stratifying breast cancer patients according to phenotypic, genotypic and/or clinical parameters into subsets, and then identifying markers that discriminate diagnosis and/or prognosis in each subsets.

5.1.1 IDENTIFICATION OF CONDITION SUBSETS

In one embodiment, the subsets are distinguished by phenotypic, genotypic, and/or clinical characteristics of breast cancer. In this embodiment, groups of individuals are

classified according to one or more phenotypic, genotypic, or clinical characteristics of breast cancer into different patient subsets. At any step in the process of subdividing a patient population into patient subsets, the expression level of one or more genes may be determined in order to identify whether a prognosis-informative set of genes may be identified for the particular patient subset. If a gene set is identified for a subset of patients, but is not as informative as desired, the patient subset may be further divided and a new geneset identified. These subsets may be further subdivided. For example, a group of individuals affected by breast cancer may be classified first on the basis of a phenotypic, genotypic or clinical characteristic A into subsets S1 and S2. The levels of expression of a plurality of genes are then determined in tumor samples taken from individuals that fall within subsets S1 or S2 in order to identify sets of genes informative for prognosis within these subsets. Subsets S1 and S2 may then each be subdivided into two or more subsets based on other phenotypic, genotypic or clinical characteristics. The basis for subdivision, if performed, need not be the same for S1 and S2. For example, in various embodiments, S1 is not subdivided, while S2 is subdivided on the basis of characteristic B; or S1 is subdivided based on characteristic B while S2 is not subdivided; or S1 and S2 are both subdivided on the basis of characteristic B; or S1 is subdivided based on characteristic B, while S2 is subdivided according to characteristic C; and so on. For a particular decision matrix leading to a plurality of patient subsets, the preferred outcome is a prognosis-informative set of genes for each patient subset. Different decision matrices may lead to different patient subsets, which, in turn, may result in different sets of prognosis-informative genes.

In the specific example, a plurality of phenotypic, genotypic or clinical indications are used to classify a breast cancer patient as being a member of one of a plurality of patient subsets, wherein the subsets are medically, biochemically or genetically relevant to breast cancer. For example, a group of patients may be classified into patient subsets based on criteria including, but not limited to, estrogen receptor (ER) status, type of tumor (*i.e.*, *BRCA1*-type or sporadic), lymph node status, grade of cancer, invasiveness of the tumor, or age. "BRCA1-type" indicates that the *BRCA1* mutation is present. In each classification step, a group of cancer patients may be classified into only two classes, for example, ER+ or ER⁻, or into three or more subsets (for example, by tumor grade), depending upon the characteristic used to determine the subsets. As used herein, "ER+" indicates that the estrogen receptor is expressed at some level; for example, it may indicate that the estrogen receptor is detectably expressed, or may indicate that more than 10% of cells may be

histologically stained for the receptor, etc. Conversely, "ER-" indicates that the estrogen receptor is expressed at a reduced level or not at all; for example, it may indicate that the receptor is not detectably expressed, or that 10% or less of cells may be histologically stained for the receptor, etc. Marker gene sets optimized for each phenotypic class are preferably determined after the subsets are established. Where informative markers for a particular patient subset, distinguished from another subset by a particular characteristic of the condition of interest, cannot be determined, the subset may be further divided by another characteristic of the condition to create a plurality of second patient subsets, whereupon genes informative for these second patient subsets may be identified.

FIG. 1 depicts the process, described in the Examples, of subdivision of a collection of breast cancer patients according to phenotypic and genotypic characteristics relevant to breast cancer, in preparation for identification of genes informative for prognosis. A collection of breast cancer tumor samples was first subdivided by estrogen receptor status. ER status was chosen because the presence or absence of the estrogen receptor greatly influences the expression of other genes. In the ER+ patient subset, it was noted that patients appeared to be bimodally distributed by ER level vs. age; that is, ER level dependence upon age tended to fall within two classes, as separated by the solid line in FIG. 2A. This bimodality was used to further subdivide ER+ individuals into "ER+, ER/AGE high" individuals and "ER+, ER/AGE low" individuals. A set of informative genes was identified for the ER+, ER/AGE high patient subset. An informative set was not identified for the ER+, ER/AGE low subset, however, so the subset of patients was further divided into LN+ and LN- individuals.

5.1.2 IDENTIFICATION OF MARKER SETS INFORMATIVE FOR PATIENT SUBSETS

Once a patient subset is identified, markers, such as genes, informative for a particular outcome, such as prognosis, may be identified. The methods make use of measured cellular constituent profiles, e.g., expression profiles of a plurality of genes (e.g., measurements of abundance levels of the corresponding gene products), in tumor samples from a plurality of patients in the patient subset whose prognosis outcomes are known. The prognosis outcomes can be the prognosis at a predetermined time after initial diagnosis. The predetermined time can be any convenient time period, e.g., 2, 3, 4, or 5 years. Prognosis markers can be obtained by identifying genes whose expression levels correlate with prognosis outcome, e.g.,

genes whose expression levels in good prognosis patients group are significantly different from those in poor prognosis patients. In preferred embodiments, the tumor samples from the plurality of patients are separated into a good prognosis group and a poor prognosis group for the predetermined time period. Genes whose expression levels exhibit differences between the good and poor prognosis groups to at least a predetermined level are selected as the genes whose expression levels correlate with patient prognosis. This section describes embodiments which employ genes and gene-derived nucleic acids as markers. However, it will be understood by a person skilled in the art that proteins or other cellular constituents may also be used as markers.

- 10 In a preferred embodiment, the expression profile is a differential expression profile. Each measurement in the profile is a differential expression level of a marker in a breast tumor sample versus that in a reference sample (also termed a standard or control sample). In one embodiment, the reference sample comprises polynucleotide molecules, derived from one or more samples from a plurality of normal individuals. For example, the normal
- 15 individuals may be persons not afflicted with breast cancer. The standard or control may alternatively comprise polynucleotide molecules, derived from one or more samples derived from individuals having a different form or stage of breast cancer; a different disease or different condition, or individuals exposed or subjected to a different condition, than the individual from which the sample of interest was obtained. The reference or control may be a
- 20 sample, or set of samples, taken from the individual at an earlier time, for example, to assess the progression of a condition, or the response to a course of therapy.

- In a preferred embodiment, the standard or control is a pool of target polynucleotide molecules derived from a plurality of different individuals. However, where protein levels, or the levels of any other relevant biomolecule, are to be compared, the pool may be a pool of
- 25 proteins or the relevant biomolecule. In a preferred embodiment in the context of breast cancer, the pool comprises samples taken from a number of individuals having sporadic-type tumors.

- In another preferred embodiment, the pool comprises an artificially-generated population of nucleic acids designed to approximate the level of nucleic acid derived from each marker found in a pool of marker-derived nucleic acids derived from tumor samples. In
- 30 another embodiment, the pool, also called a "mathematical sample pool," is represented by a set of expression values, rather than a set of physical polynucleotides; the level of expression

of relevant markers in a sample from an individual with a condition, such as a disease, is compared to values representing control levels of expression for the same markers in the mathematical sample pool. Such a control may be a set of values stored on a computer. Such artificial or mathematical controls may be constructed for any condition of interest.

- 5 In another embodiment, the reference sample is derived from a normal breast cell line or a breast cancer cell line. In a preferred embodiment, the reference sample comprises samples taken from individuals within a specific patient subset, *e.g.*, "ER+, ER/AGE high" individuals, wherein each of said individuals has a good prognosis, or each of said individuals has a poor prognosis. Of course, where, for example, expressed proteins are used as markers, 10 the proteins are obtained from the individual's sample, and the standard or control can be a pool of proteins from a number of normal individuals, or from a number of individuals having a particular state of a condition, such as a pool of samples from individuals having a particular prognosis of breast cancer.

- The comparison may be accomplished by any means known in the art. For example, 15 expression levels of various markers may be assessed by separation of target polynucleotide molecules (*e.g.*, RNA or cDNA) derived from the markers in agarose or polyacrylamide gels, followed by hybridization with marker-specific oligonucleotide probes. Alternatively, the comparison may be accomplished by the labeling of target polynucleotide molecules followed by separation on a sequencing gel. Polynucleotide samples are placed on the gel 20 such that patient and control or standard polynucleotides are in adjacent lanes. Comparison of expression levels is accomplished visually or by means of densitometer. In a preferred embodiment, the expression of all markers is assessed simultaneously by hybridization to a microarray. In each approach, markers meeting certain criteria are identified as informative for the prognosis of breast cancer.

- 25 In one embodiment, genes are first screened based on significant variation in expression in a set of breast cancer tumor samples as compared to a standard or control sample. Genes may be screened, for example, by determining whether they show significant variation in at least some samples among the set of samples. Genes that do not show significant variation in at least some samples in the set of samples are presumed not to be 30 informative, and are discarded from further consideration. Genes showing significant variation in at least some samples in the sample set are retained as candidate informative genes. The degree of variation in expression of a gene may be estimated by determining a

difference or ratio of the expression of the gene in a sample and a control. The difference or ratio of expression may be further transformed, *e.g.*, by a linear or log transformation.

Selection of candidate markers may be made based upon either significant up- or down-regulation of the gene in at least some samples in the set or based on the statistical
 5 significance (*e.g.*, the *p*-value) of the variation in expression of the gene. Preferably, both selection criteria are used. Thus, in one embodiment of the present invention, genes showing both a more than two-fold change (increase or decrease) in expression as compared to a standard in at least three samples, and a *p*-value of variation in expression of the gene in the set of tumor samples as compared to the standard sample is no more than 0.01 (*i.e.*, is
 10 statistically significant) are selected as candidate genes associated with prognosis of breast cancer in a patient subset.

In the present invention, a "good prognosis" predicts non-occurrence of metastases within a predetermined period of time, *e.g.*, 1, 2, 3, 4, 5 or more years, after initial diagnosis, whereas a "poor prognosis" predicts occurrence of metastasis within that period. In a specific
 15 embodiment, a "good prognosis" predicts non-occurrence of metastases within 5 years after initial diagnosis, and a "poor prognosis" predicts occurrence of metastasis within that period.

Expression profiles comprising a plurality of different genes in a plurality of *N* breast cancer tumor samples can be used to identify markers that correlate with, and therefore are useful for discriminating, different clinical categories. In a specific embodiment, a
 20 correlation coefficient ρ between a vector \vec{c} representing clinical categories or clinical parameters, *e.g.*, a good or poor prognosis, in the *N* tumor samples and a vector \vec{r} representing the measured expression levels of a gene in the *N* tumor samples is used as a measure of the correlation between the expression level of the gene and clinical category. The expression levels can be a measured abundance level of a transcript of the gene, or any
 25 transformation of the measured abundance, *e.g.*, a logarithmic or a log ratio. Specifically, the correlation coefficient may be calculated as:

$$\rho = (\vec{c} \cdot \vec{r}) / (\|\vec{c}\| \cdot \|\vec{r}\|) \quad (1)$$

Markers for which the coefficient of correlation exceeds a cutoff are identified as prognosis-informative markers specific for a particular clinical category, *e.g.*, a good
 30 prognosis, within a given patient subset. Such a cutoff or threshold may correspond to a certain significance of the set of obtained discriminating genes. The threshold may also be

selected based on the number of samples used. For example, a threshold can be calculated as $3 \times 1/\sqrt{n-3}$, where $1/\sqrt{n-3}$ is the distribution width and n = the number of samples. In a specific embodiment, markers are chosen if the correlation coefficient is greater than about 0.3 or less than about -0.3.

- 5 Next, the significance of the set of marker genes can be evaluated. The significance may be calculated by any appropriate statistical method. In a specific example, a Monte-Carlo technique is used to randomize the association between the expression profiles of the plurality of patients and the clinical categories to generate a set of randomized data. The same marker selection procedure as used to select the marker set is applied to the randomized data to obtain a control marker set. A plurality of such runs can be performed to generate a probability distribution of the number of genes in control marker sets. In a preferred embodiment, 10,000 such runs are performed. From the probability distribution, the probability of finding a marker set consisting of a given number of markers when no correlation between the expression levels and phenotype is expected (i.e., based randomized data) can be determined. The significance of the marker set obtained from the real data can be evaluated based on the number of markers in the marker set by comparing to the probability of obtaining a control marker set consisting of the same number of markers using the randomized data. In one embodiment, if the probability of obtaining a control marker set consisting of the same number of markers using the randomized data is below a given probability threshold, the marker set is said to be significant.

- 25 Once a marker set is identified, the markers may be rank-ordered in order of correlation or significance of discrimination. One means of rank ordering is by the amplitude of correlation between the change in gene expression of the marker and the specific condition being discriminated. Another, preferred, means is to use a statistical metric. In a specific embodiment, the metric is a t-test-like statistic:

$$t = \frac{(\langle x_1 \rangle - \langle x_2 \rangle)}{\sqrt{[\sigma_1^2(n_1 - 1) + \sigma_2^2(n_2 - 1)] / (n_1 + n_2 - 1) / (1/n_1 + 1/n_2)}} \quad (2)$$

In this equation, $\langle x_1 \rangle$ is the error-weighted average of the log ratio of transcript expression measurements within a first clinical group (e.g., good prognosis), $\langle x_2 \rangle$ is the error-weighted average of log ratio within a second, related clinical group (e.g., poor prognosis), σ_1 is the

variance of the log ratio within the first clinical group (*e.g.*, good prognosis), n_1 is the number of samples for which valid measurements of log ratios are available, σ_2 is the variance of log ratio within the second clinical group (*e.g.*, poor prognosis), and n_2 is the number of samples for which valid measurements of log ratios are available. The t -value represents the variance-compensated difference between two means. The rank-ordered marker set may be used to optimize the number of markers in the set used for discrimination.

A set of genes for prognosis of breast cancer can also be identified using an iterative approach. This is accomplished generally in a "leave one out" method as follows. In a first run, a subset, for example five, of the markers from the top of the ranked list is used to generate a template, where out of N samples, $N-1$ are used to generate the template, and the status of the remaining sample is predicted. This process is repeated for every sample until every one of the N samples is predicted once. In a second run, one or more additional markers, for example five additional markers, are added, so that a template is now generated from 10 markers, and the outcome of the remaining sample is predicted. This process is repeated until the entire set of markers is used to generate the template. For each of the runs, type 1 error (false negative) and type 2 errors (false positive) are counted. The set of top-ranked markers that corresponds to lowest type 1 error rate, or type 2 error rate, or preferably the total of type 1 and type 2 error rate is selected.

For prognostic markers, validation of the marker set may be accomplished by an additional statistic, a survival model. This statistic generates the probability of tumor distant metastases as a function of time since initial diagnosis. A number of models may be used, including Weibull, normal, log-normal, log logistic, log-exponential, or log-Rayleigh (Chapter 12 "Life Testing", S-PLUS 2000 GUIDE TO STATISTICS, Vol. 2, p. 368 (2000)). For the "normal" model, the probability of distant metastases P at time t is calculated as

$$P = \alpha \times \exp\left(-t^2 / \tau^2\right) \quad (3)$$

where α is fixed and equal to 1, and τ is a parameter to be fitted and measures the "expected lifetime".

It is preferable that the above marker identification process be iterated one or more times by excluding one or more samples from the marker selection or ranking (*i.e.*, from the

calculation of correlation). Those samples being excluded are the ones that can not be predicted correctly from the previous iteration. Preferably, those samples excluded from marker selection in this iteration process are included in the classifier performance evaluation, to avoid overstating the performance.

5 5.1.3 CLASSIFIER GENESETS FOR PATIENT SUBSETS

This section provides exemplary sets of markers useful for the prognosis of breast cancer. The markers were identified according to the above methods in specific subsets of individuals with breast cancer. Generally, the marker sets can be used for prognosis of breast cancer patients that fall into five phenotypic categories based on criteria relevant to breast cancer prognosis, including estrogen receptor (ER) status, lymph node status, type of mutation(s) (*i.e.*, *BRCA1*-type or sporadic) and age at diagnosis. More specifically, patients, and tumors from which samples were taken, were classified as ER⁻, sporadic (*i.e.*, being both estrogen receptor negative and having a non-*BRCA1*-type tumor); ER⁻, *BRCA1* (*i.e.*, being both estrogen receptor negative and having a *BRCA1*-type tumor); ER⁺, ER/AGE high (*i.e.*, estrogen receptor positive with a high ratio of the log (ratio) of estrogen receptor gene expression to age); ER⁺, ER/AGE low, LN⁺ (*i.e.*, estrogen receptor positive with a low ratio of the log (ratio) of estrogen receptor gene expression to age, lymph node positive); and ER⁺, ER/AGE low, LN⁻ (*i.e.*, estrogen receptor positive with a low ratio of the log (ratio) of estrogen receptor gene expression to age, lymph node negative). The rationale for subdivision of the original patient set into these five subsets is detailed in the Examples (Section 6). The marker sets useful for each of the subsets above are provided in Tables 1-5, respectively.

Table 1: Geneset of 20 markers used to classify ER⁻, sporadic individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
AF055033	IGFBP5	-2.12	0.88	0.54	insulin-like growth factor binding protein 5	Growth factor binding, Glycoprotein, Signal, 3D-structure
NM_000599	IGFBP5	-3.41	0.43	0.53	insulin-like growth factor binding protein 5	Growth factor binding, Glycoprotein, Signal, 3D-structure
L27560	IGFBP5	-4.55	0	0.52	EST	Hypothetical protein
AF052162	FLJ12443	-0.27	1.6	0.52	EST	Hypothetical protein
NM_001456	FLNA	-0.61	2.47	0.52	filamin A, alpha (actin binding protein 280)	Hypothetical protein, Actin-binding, Phosphorylation, Repeat,

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
						Polymorphism, Disease mutation
NM_002205	ITGA5	-0.37	2.08	0.49	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	Integrin, Cell adhesion, Receptor, Glycoprotein, Transmembrane, Signal, Calcium, Repeat
NM_013261	PPARGC1	0.09	1.54	0.47	peroxisome proliferative activated receptor, gamma, coactivator 1	
NM_001605	AARS	0.39	2.36	0.51	alanyl-tRNA synthetase	Aminoacyl-tRNA synthetase, Protein biosynthesis, Ligase, ATP-binding
X87949	HSPA5	-0.03	2.03	0.49	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	ATP-binding, Hypothetical protein, Endoplasmic reticulum, Signal
Contig50950_RC	NGEF	-1.17	3.2	0.52	neuronal guanine nucleotide exchange factor	
NM_005689	ABCB6	-0.51	2.26	0.48	ATP-binding cassette, sub-family B (MDR/TAP), member 6	ATP-binding, Transport, Transmembrane, Mitochondrion, Inner membrane, Transit peptide, Hypothetical protein
NM_004577	PSPH	-0.56	3.05	0.51	phosphoserine phosphatase	Hydrolase, Serine biosynthesis, Magnesium, Phosphorylation
NM_003832	PSPHL	-2.08	2.18	0.5	phosphoserine phosphatase-like	
NM_002422	MMP3	-0.96	2.54	0.5	matrix metalloproteinase 3 (stromelysin 1, progelatinase)	Hydrolase, Metalloprotease, Glycoprotein, Zinc, Zymogen, Calcium, Collagen degradation, Extracellular matrix, Signal, Polymorphism, 3D-structure
Contig37562_RC		-3.42	-6.02	-0.59	ESTs	
NM_018465	MDS030	-0.82	-3.28	-0.58	uncharacterized hematopoietic stem/progenitor cells protein MDS030	Hypothetical protein
Contig54661_RC		-0.79	-2.08	-0.54	ESTs	
AB032969	KIAA1143	-0.6	-2.85	-0.53	KIAA1143 protein	Hypothetical protein
Contig55353_RC	KIAA1915	-0.27	-1.82	-0.47	KIAA1915 protein	Hypothetical protein
NM_006213	CSTA	2.11	-3.4	-0.49	cystatin A (stefin A)	Thiol protease inhibitor, 3D-structure

Table 2. Geneset of 10 markers used to classify ER⁻, BRCA1 individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_list
AF005487		6.08	0.5	-0.79	HLA-DRB6	Homo sapiens MHC class II antigen (DRB6)	MHC

						mRNA, HLA-DRB6*0201 allele, sequence.	
Contig50728_RC		4.02	0.25	-0.77		ESTs, Weakly similar to S26650 DNA-binding protein 5 - human [H.sapiens]	
Contig53598_RC		8.41	3.26	-0.77	FLJ11413	hypothetical protein FLJ11413	Hypothetical protein
NM_002888	RARR ES1	6.9	0.05	-0.87	RARRES1	retinoic acid receptor responder (tazarotene induced) 1	Receptor, Transmembrane, Signal-anchor
NM_005218	DEFB1	5.14	-3.02	-0.81	DEFB1	defensin, beta 1	Antibiotic, Signal, 3D-structure
U17077	BENE	2.72	-1.72	-0.77	BENE	BENE protein	Transmembrane
Contig14683_RC		1.29	-2.31	-0.74		ESTs	
Contig53641_RC		-3.29	4.23	0.75	MAGE-E1	MAGE-E1 protein	Hypothetical protein
Contig56678_RC		-6.7	-9.73	-0.82		ESTs, Highly similar to THYA_HUMAN Prothymosin alpha [H.sapiens]	
NM_005461	KRML	0.88	-3.38	-0.75	MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	Transcription regulation, Repressor, DNA-binding, Nuclear protein, Hypothetical protein

Table 3. Geneset of 50 markers used to classify ER+, ER/AGE high individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
NM_003600	STK15	-2.93	2.08	0.8	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase
NM_003158	STK6	-1.57	1.42	0.78	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase
NM_007019	UBCH10	-2.98	2.62	0.81	ubiquitin-conjugating enzyme E2C	Hypothetical protein, Ubi conjugation pathway, Ligase, Multigene family, Mitosis, Cell cycle, Cell division
NM_013277	ID-GAP	-2.43	2.43	0.77	Rac GTPase activating protein 1	Hypothetical protein
NM_004336	BUB1	-2.04	1.39	0.77	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)	Transferase, Serine/threonine-protein kinase, ATP-binding, Cell cycle, Nuclear protein, Mitosis, Phosphorylation, Polymorphism
NM_006607	PTTG2	-1.71	1.49	0.72	pituitary tumor-transforming 2	
AK001166	FLJ11252	-1.33	0.99	0.71	hypothetical protein FLJ11252	Hypothetical protein
NM_004701	CCNB2	-4.62	2.01	0.81	cyclin B2	Cyclin, Cell cycle, Cell division, Mitosis
Contig57584_RC		-3.68	2.04	0.78	likely ortholog of mouse gene rich cluster, C8 gene	
NM_006845	KNSL6	-4.13	1.05	0.73	kinesin-like 6 (mitotic)	Hypothetical protein,

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
					centromere-associated kinesin)	Motor protein, Microtubules, ATP-binding, Coiled coil, Nuclear protein
Contig38901_RC		-3.08	1.15	0.75	hypothetical protein MGC45866	Hypothetical protein
NM_018410	DKFZp762E1312	-4.38	1.49	0.75	hypothetical protein DKFZp762E1312	Hypothetical protein
NM_003981	PRC1	-3.52	2.17	0.78	protein regulator of cytokinesis 1	
NM_001809	CENPA	-5.04	0.98	0.75	centromere protein A, 17kDa	Hypothetical protein, Chromosomal protein, Nuclear protein, DNA-binding, Centromere, Antigen
NM_003504	CDC45L	-2.67	1.22	0.73	CDC45 cell division cycle 45-like (S. cerevisiae)	DNA replication, Cell cycle, Nuclear protein, Cell division
Contig41413_RC		-5.43	2.15	0.74	ribonucleotide reductase M2 polypeptide	Oxidoreductase, DNA replication, Iron
NM_004217	STK12	-2.17	0.73	0.72	serine/threonine kinase 12	Hypothetical protein, ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase
NM_002358	MAD2L1	-2.65	2.27	0.83	MAD2 mitotic arrest deficient-like 1 (yeast)	Cell cycle, Mitosis, Nuclear protein, 3D-structure
NM_014321	ORC6L	-2.73	1.8	0.75	origin recognition complex, subunit 6 homolog-like (yeast)	Hypothetical protein, DNA replication, Nuclear protein, DNA-binding
NM_012291	KIAA0165	-1.52	1.55	0.71	extra spindle poles like 1 (S. cerevisiae)	Hypothetical protein
NM_004203	PKMYT1	-3.64	2.2	0.7	retinoblastoma-like 2 (p130)	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase, Transcription regulation, DNA-binding, Nuclear protein, Cell cycle, Phosphorylation, Anti-oncogene
M96577	E2F1	-2.14	1.42	0.75	E2F transcription factor 1	Transcription regulation, Activator, DNA-binding, Nuclear protein, Phosphorylation, Cell cycle, Apoptosis, Polymorphism
NM_002266	KPNA2	-3.77	1.78	0.71	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Transport, Protein transport, Repeat, Nuclear protein, Polymorphism
Contig31288_RC		-2.63	0.7	0.68	ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens]	
NM_014501	E2-EPF	-1.55	1.93	0.7	ubiquitin carrier protein	Ubi conjugation pathway, Ligase, Multigene family

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
NM_001168	BIRC5	-5.76	2.01	0.78	baculoviral IAP repeat-containing 5 (survivin)	Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure, Hypothetical protein, Protease, Receptor
NM_003258	TK1	-4.57	1.38	0.71	thymidine kinase 1, soluble	Transferase, Kinase, DNA synthesis, ATP-binding
NM_001254	CDC6	-2.46	0.28	0.72	CDC6 cell division cycle 6 homolog (S. cerevisiae)	ATP-binding, Cell division
NM_004900	DJ742C19.2	-2.96	0.13	0.69	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	Hydrolase
NM_004702	CCNE2	-3.12	2.13	0.81	cyclin E2	Cell cycle, Cell division, Cyclin, Hypothetical protein, Phosphorylation, Alternative splicing, Nuclear protein
AL160131		-3.07	2.42	0.7	hypothetical protein MGC861	Hypothetical protein
NM_016359	LOC51203	-3.22	2.61	0.76	nucleolar protein ANKT	Hypothetical protein, Nuclear protein
NM_004856	KNSL5	-1.52	1.1	0.71	kinesin-like 5 (mitotic kinesin-like protein 1)	Motor protein, Cell division, Microtubules, ATP-binding, Coiled coil, Mitosis, Cell cycle, Nuclear protein
NM_000057	BLM	-1.54	0.76	0.71	Bloom syndrome	Hydrolase, Helicase, ATP-binding, DNA-binding, Nuclear protein, DNA replication, Disease mutation
NM_018455	BM039	-2.44	1.18	0.7	uncharacterized bone marrow protein BM039	
NM_002106	H2AFZ	-2.49	1.53	0.72	H2A histone family, member Z	Chromosomal protein, Nucleosome core, Nuclear protein, DNA-binding, Multigene family
Contig64688		-2.68	3.1	0.73	hypothetical protein FLJ23468	Hypothetical protein
Contig44289_RC		-1.65	1.6	0.67	ESTs	
Contig28552_RC		-1.37	1.53	0.68	diaphanous homolog 3 (Drosophila)	Hypothetical protein, Coiled coil, Repeat, Alternative splicing
Contig46218_RC		-1.31	1.56	0.68	ESTs, Weakly similar to T19201 hypothetical protein C11G6.3 - Caenorhabditis elegans [C. elegans]	
Contig28947_RC		-1.3	0.98	0.67	cell division cycle 25A	Hypothetical protein, Cell division, Mitosis, Hydrolase, Alternative splicing, Multigene family, 3D-structure
NM_016095	LOC5165	-1.4	2.13	0.67	HSPC037 protein	Hypothetical protein

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
	9					
NM_003090	SNRPA1	-3.26	0.95	0.7	small nuclear ribonucleoprotein polypeptide A'	Hypothetical protein, Nuclear protein, RNA- binding, Ribonucleoprotein, Leucine-rich repeat, Repeat, 3D-structure
NM_002811	PSMD7	-2.48	1.89	0.7	proteasome (prosome, macropain) 26S subunit, non- ATPase, 7 (Mov34 homolog)	Proteasome
Contig38288_RC		-2.34	0.97	0.67	hypothetical protein DKFZp762A2013	Hypothetical protein
NM_003406	YVHAZ	-1.5	2.79	0.68	tyrosine 3- monooxygenase/tryp- tophan 5- monooxygenase activation protein, zeta polypeptide	Brain, Neurone, Phosphorylation, Acetylation, Multigene family, 3D-structure
AL137540	NTN4	2.13	-4.61	-0.69	netrin 4	Hypothetical protein, Laminin EGF-like domain, Signal
AL049367		1.9	-3.2	-0.68	EST	Transducer, Prenylation, Lipoprotein, Multigene family, Acetylation
NM_013409	FST	1.04	-5.78	-0.69	folistatin	Glycoprotein, Repeat, Signal, Alternative splicing
NM_000060	BTD	3.1	-1.45	-0.67	biotinidase	Hydrolase, Glycoprotein, Signal, Disease mutation

Table 4. Geneset of 50 markers used to classify ER+, ER/AGE low, LN+ individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
NM_006417	MTAP44	-1.5	3	0.69	Fc fragment of IgG, low affinity Iib, receptor for (CD32)	Hydrolase, Hypothetical protein, immunoglobulin domain, IgG-binding protein, Receptor, Transmembrane, Glycoprotein, Signal, Repeat, Multigene family, Polymorphism, NAD, One-carbon metabolism, Serine protease, Zymogen, Protease, Alternative splicing, Chromosomal translocation, Proto- oncogene, Galaptin, Lectin, Antigen
NM_006820	GS3686	-4.3	4.06	0.69	chromosome 1 open reading frame 29	Hypothetical protein

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
NM_001548	IFIT1	-3.4	4.27	0.71	Interferon-induced protein with tetratricopeptide repeats 1	Repeat, TPR repeat, Interferon induction
Contig41538_RC		-2.5	3.16	0.68	ESTs, Moderately similar to hypothetical protein FLJ20489 [<i>Homo sapiens</i>]	
NM_016816	OAS1	-1.7	3.29	0.75	2',5'-oligoadenylate synthetase 1, 40/46kDa	RNA-binding, Transferase, Nucleotidyltransferase, Interferon induction, Alternative splicing
Contig51660_RC		-2.1	2.65	0.66	28kD interferon responsive protein	Transmembrane
Contig43645_RC		-4.8	1.44	0.63	<i>Homo sapiens</i> , clone IMAGE:4428577, mRNA, partial cds	Hypothetical protein
AF026941		-4.6	2.71	0.63	EST, Weakly similar to 2004399A chromosomal protein [<i>Homo sapiens</i>]	Hypothetical protein
NM_007315	STAT1	-3.5	1.8	0.59	signal transducer and activator of transcription 1, 91kDa	Transcription regulation, DNA-binding, Nuclear protein, Phosphorylation, SH2 domain, Alternative splicing, 3D-structure
NM_002038	G1P3	-4.1	5.64	0.79	interferon, alpha-inducible protein (clone IFI-6-16)	Interferon induction, Transmembrane, Signal, Alternative splicing
NM_005101	ISG15	-5.6	5.34	0.77	interferon-stimulated protein, 15 kDa	Interferon induction, Repeat
NM_002462	MX1	-6.1	0.83	0.56	myxovirus (influenza virus) resistance 1, Interferon-inducible protein p78 (mouse)	Hypothetical protein, Interferon induction, GTP-binding, Multigene family, Antiviral
NM_005532	IFI27	-5.8	2.81	0.59	interferon, alpha-inducible protein 27	Interferon induction, Transmembrane
NM_002346	LY6E	-2.1	3.58	0.75	lymphocyte antigen 6 complex, locus E	Signal, Antigen, Multigene family, Membrane, GPI-anchor
NM_016817	OAS2	-3.6	1.89	0.59	2'-5'-oligoadenylate synthetase 2, 69/71kDa	RNA-binding, Transferase, Nucleotidyltransferase, Repeat, Interferon induction, Alternative splicing, Myristate
Contig44909_RC		-2.3	1.13	0.55	hypothetical protein BC012330	Hypothetical protein
NM_017414	USP18	-4.1	3.37	0.72	ubiquitin specific protease 18	Ubl conjugation pathway, Hydrolase, Thiol protease, Multigene family
NM_004029	IRF7	-2.4	3.67	0.66	interferon regulatory factor 7	Collagen, Transcription regulation, DNA-binding, Nuclear protein, Activator, Alternative splicing

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
NM_004335	BST2	-3.2	3.22	0.57	bone marrow stromal cell antigen 2	Transmembrane, Glycoprotein, Signal- anchor, Polymorphism
NM_002759	PRKR	-2.4	1.8	0.58	protein kinase, interferon-inducible double stranded RNA dependent	Transferase, Serine/threonine-protein kinase, ATP-binding, Repeat, Phosphorylation, Interferon induction, RNA- binding, 3D-structure
NM_006332	IFI30	-3.8	2.65	0.64	interferon, gamma- inducible protein 30	Oxidoreductase, Interferon induction, Glycoprotein, Lysosome, Signal, Hypothetical protein
NM_009587	LGALS9	-3.2	2.08	0.6	lectin, galactoside- binding, soluble, 9 (galectin 9)	Galapin, Lectin, Repeat, Alternative splicing
NM_003641	IFITM1	-2.4	5.54	0.63	interferon induced transmembrane protein 1 (9-27)	Interferon induction, Transmembrane
NM_017523	HSXIAPA F1	-1	2.84	0.7	XIAP associated factor-1	Hypothetical protein
NM_014314	RIG-I	-1.3	3.55	0.62	RNA helicase	ATP-binding, Helicase, Hydrolase, Hypothetical protein
Contig47563_RC		-2.2	3.11	0.56	ESTs	
A1497657_RC		-4.4	5.61	0.74	guanine nucleotide binding protein 4	Transducer, Prenylation, Lipoprotein, Multigene family
NM_000735	CGA	-4.3	2.5	0.58	glycoprotein hormones, alpha polypeptide	Hormone, Glycoprotein, Signal, 3D-structure
NM_004988	MAGEA1	-1.4	6.31	0.64	melanoma antigen, family A, 1 (directs expression of antigen MZ2-E)	Antigen, Multigene family, Polymorphism, Tumor antigen
Contig54242_RC		-1.2	4.1	0.65	chromosome 17 open reading frame 26	Hypothetical protein
NM_004710	SYNGR2	-1.4	3.01	0.54	synaptogyrin 2	Transmembrane
NM_001168	BIRC5	-3.7	3.39	0.64	baculoviral IAP repeat-containing 5 (survivin)	Hypothetical protein, Protease, Receptor, Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure
Contig41413_RC		-4.4	2.61	0.57	ribonucleotide reductase M2 polypeptide	Oxidoreductase, DNA replication, Iron
NM_004203	PKMYT1	-3.4	3.79	0.6	retinoblastoma-like 2 (p130)	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase, Transcription regulation, DNA-binding, Nuclear protein, Cell cycle, Phosphorylation, Anti- oncogene

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
Contig48913_RC		-3.1	1.72	0.55	<i>Homo sapiens</i> , Similar to hypothetical protein PRO1722, clone MGC:15692 IMAGE:3351479, mRNA, complete cds	
NM_005804	DDXL	-2.5	1.42	0.58	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 39	ATP-binding, Helicase, Hydrolase, Hypothetical protein
NM_016359	LOC51203	-1.7	3.6	0.57	nucleolar protein ANKT	Hypothetical protein, Nuclear protein
NM_001645	APOC1	-2.9	3.43	0.58	apolipoprotein C-I	Plasma, Lipid transport, VLDL, Signal, 3D-structure, Polymorphism
Contig37895_RC		-2	2.05	0.55	ESTs	
NM_005749	TOB1	-1.3	4.96	0.59	transducer of ERBB2, 1	Phosphorylation
NM_000269	NME1	-1.3	2.98	0.55	non-metastatic cells 1, protein (NM23A) expressed in	Transferase, Kinase, ATP-binding, Nuclear protein, Anti-oncogene, Disease mutation
NM_014462	LSM1	-1	4.5	0.57	Lsm1 protein	Nuclear protein, Ribonucleoprotein, mRNA splicing, mRNA processing, RNA-binding
Contig31221_RC		-1.4	3.83	0.56	HTPAP protein	
NM_005326	HAGH	-1.9	4.29	0.57	hydroxyacyl glutathione hydrolase	Hydrolase, Zinc, 3D-structure
Contig42342_RC		0.78	-3.2	-0.6	<i>Homo sapiens</i> cDNA FLJ39417 fis, clone PLACE6016942	Hypothetical protein
AL137540	NTN4	2.24	-3.9	-0.6	netrin 4	Laminin EGF-like domain, Signal, Hypothetical protein
Contig40434_RC		1.64	-5.6	-0.6	wingless-type MMTV integration site family, member 5A	Developmental protein, Glycoprotein, Signal
Contig1632_RC		1.03	-3.9	-0.6	hypothetical protein MGC17921	Hypothetical protein
NM_014246	CELSR1	0.95	-4.6	-0.6	cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, <i>Drosophila</i>)	G-protein coupled receptor, Transmembrane, Glycoprotein, EGF-like domain, Calcium-binding, Laminin EGF-like domain, Repeat, Developmental protein, Hydroxylation, Signal, Alternative splicing, Hypothetical protein
NM_005139	ANXA3	1.26	-6.2	-0.6	annexin A3	Annexin, Calcium/phospholipid-binding, Repeat, Phospholipase A2 inhibitor, 3D-structure, Polymorphism

Table 5. Geneset of 65 markers used to classify ER+, ER/AGE low, LN⁻ individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
M55914	MPB1	-2.82	1.25	0.5	ENO1	enolase 1, (alpha)	DNA-binding, Transcription regulation, Repressor, Nuclear protein, Lyase, Glycolysis, Magnesium, Multigene family, Hypothetical protein
NM_005945	MPB1	-3.06	1.19	0.49	ENO1	Homo sapiens enolase 1, (alpha) (ENO1), mRNA.	Glycolysis, Hypothetical protein, Lyase, Magnesium, DNA-binding, Transcription regulation, Repressor, Nuclear protein, Multigene family
NM_001428	ENO1	-2.53	1.18	0.46	ENO1	enolase 1, (alpha)	DNA-binding, Transcription regulation, Repressor, Nuclear protein, Lyase, Glycolysis, Magnesium, Multigene family, Hypothetical protein
NM_001216	CA9	-4.72	1.49	0.6	CA9	carbonic anhydrase IX	Lyase, Zinc, Transmembrane, Glycoprotein, Antigen, Signal, Nuclear protein, Polymorphism
NM_001124	ADM	-5.68	2.99	0.56	ADM	adrenomedullin	Hormone, Amidation, Cleavage on pair of basic residues, Signal
NM_000584	IL8	-2.45	2.04	0.54	IL8	interleukin 8	Cytokine, Chemotaxis, Inflammatory response, Signal, Alternative splicing, 3D-structure
D25328	PFKP	-4.19	3.29	0.56	PFKP	Phosphofructo- kinase, platelet	Kinase, Transferase, Glycolysis, Repeat, Allosteric enzyme, Phosphorylation, Magnesium, Multigene family
NM_006096	NDRG1	-5.45	5.97	0.77	NDRG1	N-myc downstream regulated gene 1	Hypothetical protein, Nuclear protein, Repeat
NM_004994	MMP9	-5.53	1.07	0.49	MMP9	matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	Hydrolase, Metalloprotease, Glycoprotein, Zinc, Zymogen, Calcium, Collagen degradation, Extracellular matrix, Repeat, Signal,

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
							Polymorphism, 3D- structure
NM_003311	TSSC3	-4.57	5.58	0.68	TSSC3	tumor suppressing subtransferable candidate 3	
NM_006086	TUBB4	-5.19	2.85	0.59	TUBB4	tubulin, beta, 4	G-protein coupled receptor, Transmembrane, Glycoprotein, Phosphorylation, Lipoprotein, Palmitate, Polymorphism, Hypothetical protein, GTP-binding, Receptor, Microtubules, Multigene family
NM_006115	PRAME	-4.48	2.77	0.61	PRAME	preferentially expressed antigen in melanoma	Antigen
NM_004345	CAMP	-2.02	1.37	0.49	CAMP	cathelicidin antimicrobial peptide	Antibiotic, Signal
NM_018455	BM039	-2.34	0.76	0.47	BM039	uncharacterized bone marrow protein BM039	
Contig49169_RC		-1.17	1.5	0.46	SUV39H2	suppressor of variegation 3-9 (Drosophila) homolog 2; hypothetical protein FLJ23414	Hypothetical protein, Nuclear protein
Contig45032_RC		-1.37	0.77	0.45	FLJ14813	hypothetical protein FLJ14813	Hypothetical protein, ATP-binding, Kinase, Serine/threonine- protein kinase, Transferase
NM_000917	P4HA1	-1.54	4.31	0.62	P4HA1	procollagen- proline, 2- oxoglutarate 4- dioxygenase (proline 4- hydroxylase), alpha polypeptide I	Dioxygenase, Collagen, Oxidoreductase, Iron, Vitamin C, Alternative splicing, Glycoprotein, Endoplasmic reticulum, Signal
NM_002046	GAPD	-2.51	3.42	0.6	GAPD	glyceraldehyde-3- phosphate dehydrogenase	Glycolysis, NAD, Oxidoreductase, Hypothetical protein, Multigene family
NM_000365	TPI1	-1.81	2.94	0.56	TPI1	triosephosphate isomerase 1	Fatty acid biosynthesis, Gluconeogenesis, Glycolysis, Isomerase, Pentose shunt, Disease mutation, Polymorphism, 3D-

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
NM_014364	GAPDS	-1.08	2.88	0.58	GAPDS	glyceraldehyde-3- phosphate dehydrogenase, testis-specific	structure Glycolysis, Oxidoreductase, NAD
NM_005566	LDHA	-2.01	4.01	0.59	LDHA	lactate dehydrogenase A	Oxidoreductase, NAD, Glycolysis, Multigene family, Disease mutation, Polymorphism
NM_000291	PGK1	-2.28	1.68	0.51	PGK1	phosphoglycerate kinase 1	Kinase, Transferase, Multigene family, Glycolysis, Acetylation, Disease mutation, Polymorphism, Hereditary hemolytic anemia
NM_016185	LOC511 55	-2.33	2.82	0.59	HN1	hematological and neurological expressed 1	
NM_001168	BIRC5	-4.33	2.78	0.55	BIRC5	baculoviral IAP repeat-containing 5 (survivin)	Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure, Hypothetical protein, Protease, Receptor
NM_002266	KPNA2	-3.75	1.34	0.47	KPNA2	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Transport, Protein transport, Repeat, Nuclear protein, Polymorphism
Contig31288_RC		-2.1	1.27	0.5		ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens]	
NM_000269	NME1	-2.15	3.43	0.55	NME1	non-metastatic cells 1, protein (NM23A) expressed in	Transferase, Kinase, ATP-binding, Nuclear protein, Anti- oncogene, Disease mutation
NM_003158	STK6	-1.23	1.73	0.45	STK6	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine- protein kinase, Transferase
NM_007274	HBACH	-1.83	2.73	0.51	BACH	brain acyl-CoA hydrolase	Hydrolase, Serine esterase, Repeat
Contig55188_RC		-2.36	3.28	0.47	FLJ22341	hypothetical protein FLJ22341	Hypothetical protein
NM_002061	GCLM	-1.06	1.76	0.48	GCLM	glutamate-cysteine ligase, modifier subunit	Ligase, Glutathione biosynthesis
NM_004207	SLC16A 3	-3.11	5.07	0.67	SLC16A3	solute carrier family 16 (monocarboxylic acid transporters), member 3	Transport, Symport, Transmembrane, Multigene family
NM_000582	SPP1	-5.09	5.47	0.53	SPP1	secreted phosphoprotein 1	Hypothetical protein, Glycoprotein, Sialic

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_ll st
						(osteopontin, bone sialoprotein 1, early T-lymphocyte activation 1)	acid, Biomineralization, Cell adhesion, Phosphorylation, Signal, Alternative splicing
NM_001109	ADAM8	-2.5	3.74	0.45	ADAM8	a disintegrin and metalloproteinase domain 8	Hydrolase, Metalloproteinase, Zinc, Signal, Glycoprotein, Transmembrane, Antigen
D50402	SLC11A1	-1.05	3.46	0.53	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1	Transport, Iron transport, Transmembrane, Glycoprotein, Macrophage, Polymorphism
AL080235	DKFZP586E1621	-1.23	1.96	0.51	RIS1	Ras-induced senescence 1	Hypothetical protein
Contig40552_RC		-1.26	3.96	0.54	FLJ25348	hypothetical protein FLJ25348	Hypothetical protein
Contig52490_RC		-0.64	3.33	0.61	LOC116238	hypothetical protein BC014072	
NM_006461	DEEPEST	-2.1	1.85	0.46	SPAG5	sperm associated antigen 5	Hypothetical protein
Contig56503_RC		-4.3	3.39	0.55	MGC9753	hypothetical gene MGC9753	Hypothetical protein
Contig63525		-1.91	3.34	0.5	FLJ13352	hypothetical protein FLJ13352	Hypothetical protein
NM_001909	CTSD	-0.83	4.6	0.51	CTSD	cathepsin D (lysosomal aspartyl protease)	Hydrolase, Aspartyl protease, Glycoprotein, Lysosome, Signal, Zymogen, Polymorphism, Alzheimer's disease, 3D-structure
NM_005063	SCD	-2.57	5.15	0.48	SCD	stearoyl-CoA desaturase (delta-9-desaturase)	Hypothetical protein, Endoplasmic reticulum, Fatty acid biosynthesis, Iron, Oxidoreductase, Transmembrane
NM_005165	ALDOC	-2.43	5.02	0.48	ALDOC	aldolase C, fructose-bisphosphate	Lyase, Schiff base, Glycolysis, Multigene family
NM_000363	TNNI3	-0.54	3.58	0.48	TNNI3	troponin I, cardiac	Hypothetical protein, Muscle protein, Actin-binding, Acetylation, Disease mutation, Cardiomyopathy, Receptor, Signal
AF035284		-1.63	3.28	0.47	FADS1	EST	Heme, Hypothetical protein
Contig30875_RC		-0.88	3	0.6		ESTs	

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
NM_018487	HCA112	-0.7	3.54	0.58	HCA112	hepatocellular carcinoma-associated antigen 112	Hypothetical protein
NM_001323	CST6	-1.63	3.84	0.57	CST6	cystatin E/M	Thiol protease inhibitor, Signal, Glycoprotein
NM_006516	SLC2A1	-1.66	2.22	0.46	SLC2A1	solute carrier family 2 (facilitated glucose transporter), member 1	Transmembrane, Sugar transport, Transport, Glycoprotein, Multigene family, Disease mutation
NM_007267	LAK-4P	-1.04	3.28	0.61	EVIN1	expressed in activated T/LAK lymphocytes	Hypothetical protein
NM_004710	SYNGR2	-0.84	4.81	0.56	SYNGR2	synaptogyrin 2	Transmembrane
Contig63649_RC		-1.34	6.3	0.75		ESTs, Weakly similar to 2004399A chromosomal protein [Homo sapiens] [H.sapiens]	
NM_003376	VEGF	-2.12	2.42	0.46	VEGF	vascular endothelial growth factor	Hypothetical protein, Mitogen, Angiogenesis, Growth factor, Glycoprotein, Signal, Heparin-binding, Alternative splicing, Multigene family, 3D-structure
NM_000799	EPO	-0.75	4.01	0.69	EPO	erythropoietin	Erythrocyte maturation, Glycoprotein, Hormone, Signal, Pharmaceutical, 3D-structure
NM_006014	DXS9879E	-1.85	3.44	0.54	DXS9879E	DNA segment on chromosome X (unique) 9879 expressed sequence	
NM_007183	PKP3	-0.91	4.14	0.48	PKP3	plakophilin 3	Cell adhesion, Cytoskeleton, Structural protein, Nuclear protein, Repeat
D13642	SF3B3	-0.65	2.28	0.48	SF3B3	splicing factor 3b, subunit 3, 130kDa	Hypothetical protein, Spliceosome, mRNA processing, mRNA splicing, Nuclear protein
NM_003756	EIF3S3	-1.85	2.19	0.46	EIF3S3	eukaryotic translation initiation factor 3, subunit 3 gamma, 40kDa	Initiation factor, Protein biosynthesis
Contig47096_RC		-0.41	4.52	0.54	PFKFB4	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4	Kinase, Multifunctional enzyme, Transferase, Hydrolase, ATP-binding,

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
NM_004209	SYNGR3	-0.31	3.67	0.53	SYNGR3	synaptogyrin 3	Phosphorylation, Multigene family, Transmembrane
Contig3464_RC		0.99	-5.81	-0.52		ESTs	
Contig31646_RC		1.1	-7.76	-0.5	COL14A1	collagen, type XIV, alpha 1 (undulin)	Extracellular matrix, Glycoprotein, Hypothetical protein, Collagen, Signal
Contig49388_RC		1.73	-1.75	-0.51	FLJ13322	hypothetical protein FLJ13322	Hypothetical protein
Contig41887_RC		0.37	-5.74	-0.47	LOC124220	similar to common salivary protein 1	Hypothetical protein

Table 8 Good and poor prognosis templates: mean subtracted log(intensity) values for each of the seventy markers for 44 breast cancer patients having a good prognosis (C1) or 34 breast cancer patients having a poor prognosis (C2) (see, PCT publication WO 2004/065545, published August 5, 2004, which is incorporated herein by reference in its entirety)

Marker Accession #	C1 (good prognosis template)	C2 (poor prognosis template)	SEQ ID NO
AL080059	-5.161569	0.043019	367
Contig63649_RC	-1.440895	0.966702	361
Contig46218_RC	-0.937662	0.815081	321
NM_016359	-1.49878	0.872829	253
AA555029_RC	-1.283504	0.543442	368
NM_003748	1.355486	-0.254201	369
Contig38288_RC	-1.237495	1.085461	297
NM_003862	0.981236	-1.619658	370
Contig28552_RC	-1.296043	1.067545	281
Contig32125_RC	0.855155	-0.7338	371
U82987	1.256206	-1.362807	372
AL137718	-0.55046	0.68754	373
AB037863	0.819061	-1.621057	374
NM_020188	-1.137582	0.673123	375
NM_020974	-0.463953	-5.623268	376
NM_000127	-0.618568	0.552726	377
NM_002019	-1.409168	0.547285	378
NM_002073	-1.577177	0.417352	379
NM_000436	-0.722574	0.599239	380
NM_004994	-5.561089	-2.180659	165
Contig55377_RC	0.805683	-1.16728	381
Contig35251_RC	-0.931146	0.607562	382
Contig25991	-0.720727	1.045949	383
NM_003875	-1.206839	1.163244	384
NM_006101	-0.879965	0.628296	385
NM_003882	0.529121	-0.467098	386
NM_003607	-0.959094	0.709653	387

Marker Accession #	C1 (good prognosis template)	C2 (poor prognosis template)	SEQ ID NO
AF073519	-1.451486	0.163988	388
AF052162	-1.145575	0.192391	9
NM_000849	0.944742	-1.499473	389
Contig32185_RC	-0.887643	0.688257	390
NM_016577	-2.762008	0.081637	391
Contig48328_RC	0.405401	-2.946904	392
Contig46223_RC	0.805424	-0.581849	393
NM_015984	-1.056531	0.522176	394
NM_006117	1.129928	-1.262974	395
AK000745	-2.475715	-0.013002	396
Contig40831_RC	-1.17091	0.435754	397
NM_003239	0.457773	-2.150499	398
NM_014791	-1.14862	0.383018	399
X05610	-0.768514	0.637938	400
NM_016448	-0.713264	0.632638	401
NM_018401	0.618921	-0.286778	402
NM_000788	-0.995116	0.50246	403
Contig51464_RC	-0.663538	0.765975	404
AL080079	-1.794821	0.43708	405
NM_006931	-0.846271	0.915602	406
AF257175	1.122354	-0.721924	407
NM_014321	-1.820261	0.482287	241
NM_002916	-0.966852	0.599925	408
Contig55725_RC	-2.935162	0.623397	409
Contig24252_RC	-2.004671	0.263597	410
AF201951	0.355839	-2.296556	411
NM_005915	-0.586121	0.827714	412
NM_001282	0.762645	-0.970418	413
Contig56457_RC	-0.920808	0.588269	414
NM_000599	-3.612469	-0.714313	51
NM_020386	-0.46073	0.699313	415
NM_014889	-1.678462	0.2362	416
AF055033	-2.505271	-0.07576	11
NM_006681	-0.631302	0.584119	417
NM_007203	-1.426446	0.504624	418
Contig63102_RC	0.521511	-1.266163	419
NM_003981	-2.521877	0.552669	420
Contig20217_RC	-0.363574	0.449022	421
NM_001809	-2.171301	0.328419	81
Contig2399_RC	-1.174844	0.602523	422
NM_004702	-1.560133	0.619078	155
NM_007036	-0.950633	0.34945	423
NM_018354	-1.392354	0.347831	424

5.2. DIAGNOSTIC AND PROGNOSTIC METHODS

5.2.1 SAMPLE COLLECTION

In the present invention, markers, such as target polynucleotide molecules or proteins, are extracted from a sample taken from an individual afflicted with a condition such as breast cancer. The sample may be collected in any clinically acceptable manner, but must be collected such that marker-derived polynucleotides (*i.e.*, RNA) are preserved (if gene
5 expression is to be measured) or proteins are preserved (if encoded proteins are to be measured). For example, mRNA or nucleic acids derived therefrom (*i.e.*, cDNA or amplified DNA) are preferably labeled distinguishably from standard or control polynucleotide molecules, and both are simultaneously or independently hybridized to a microarray comprising some or all of the markers or marker sets or subsets described above.

10 Alternatively, mRNA or nucleic acids derived therefrom may be labeled with the same label as the standard or control polynucleotide molecules, wherein the intensity of hybridization of each at a particular probe is compared. A sample may comprise any clinically relevant tissue sample, such as a tumor biopsy or fine needle aspirate, or a sample of bodily fluid, such as blood, plasma, serum, lymph, ascitic fluid, cystic fluid, urine or nipple exudate. The sample
15 may be taken from a human, or, in a veterinary context, from non-human animals such as ruminants, horses, swine or sheep, or from domestic companion animals such as felines and canines.

Methods for preparing total and poly(A)⁺ RNA are well known and are described generally in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.),
20 Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)) and Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994)).

RNA may be isolated from eukaryotic cells by procedures that involve lysis of the cells and denaturation of the proteins contained therein. Cells of interest include wild-type
25 cells (*i.e.*, non-cancerous), drug-exposed wild-type cells, tumor- or tumor-derived cells, modified cells, normal or tumor cell line cells, and drug-exposed modified cells. Preferably, the cells are breast cancer tumor cells.

Additional steps may be employed to remove DNA. Cell lysis may be accomplished with a nonionic detergent, followed by microcentrifugation to remove the nuclei and hence
30 the bulk of the cellular DNA. In one embodiment, RNA is extracted from cells of the various types of interest using guanidinium thiocyanate lysis followed by CsCl centrifugation to separate the RNA from DNA (Chirgwin *et al.*, *Biochemistry* 18:5294-5299 (1979)).

Poly(A)+ RNA is selected by selection with oligo-dT cellulose (*see* Sambrook *et al.*,
MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor
Laboratory, Cold Spring Harbor, New York (1989). Alternatively, separation of RNA from
DNA can be accomplished by organic extraction, for example, with hot phenol or
5 phenol/chloroform/isoamyl alcohol.

If desired, RNase inhibitors may be added to the lysis buffer. Likewise, for certain
cell types, it may be desirable to add a protein denaturation/digestion step to the protocol.

For many applications, it is desirable to preferentially enrich mRNA with respect to
other cellular RNAs, such as transfer RNA (tRNA) and ribosomal RNA (rRNA). Most
10 mRNAs contain a poly(A) tail at their 3' end. This allows them to be enriched by affinity
chromatography, for example, using oligo(dT) or poly(U) coupled to a solid support, such as
cellulose or Sephadex™ (*see* Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY,
vol. 2, Current Protocols Publishing, New York (1994). Once bound, poly(A)+ mRNA is
eluted from the affinity column using 2 mM EDTA/0.1% SDS.

15 The sample of RNA can comprise a plurality of different mRNA molecules, each
different mRNA molecule having a different nucleotide sequence. In a specific embodiment,
the mRNA molecules in the RNA sample comprise at least 5, 10, 15, 20, 25, 30, 40 or 50
different nucleotide sequences. More preferably, the mRNA molecules of the RNA sample
comprise mRNA molecules corresponding to each of the marker genes. In another specific
20 embodiment, the RNA sample is a mammalian RNA sample.

In a specific embodiment, total RNA or mRNA from cells is used in the methods of
the invention. The source of the RNA can be cells of a plant or animal, human, mammal,
primate, non-human animal, dog, cat, mouse, rat, bird, yeast, eukaryote, prokaryote, etc. In
specific embodiments, the method of the invention is used with a sample containing total
25 mRNA or total RNA from 1×10^6 cells or less. In another embodiment, proteins can be
isolated from the foregoing sources, by methods known in the art, for use in expression
analysis at the protein level.

Probes to the homologs of the marker sequences disclosed herein can be employed
preferably when non-human nucleic acid is being assayed.

The prognosis may be carried out using a cellular constituent profile of any type of molecular markers. For example, sets of proteins informative for breast cancer prognosis may be used. Such protein markers may be measured, for example, by use of gel electrophoresis, such as one-dimensional polyacrylamide gel electrophoresis, two-dimensional polyacrylamide gel electrophoresis, nondenaturing polyacrylamide gel electrophoresis; isoelectric focusing gels, etc., by use of antibody arrays, etc. Of course, the particular template(s) used to classify the individual depends upon the type(s) of cellular constituents used as markers. For example, where nucleic acids (e.g., genes or nucleic acids derived from expressed genes) are used as markers, the template comprises nucleic acids (or the level of expression or abundance thereof); where proteins are used as markers, the template comprises proteins, for example, the level or abundance of those proteins; etc.

5.2.2. USE OF PROGNOSTIC GENESETS FOR BREAST CANCER

According to the present invention, once genesets informative for a plurality of subsets of a condition are identified, an individual is classified into one of these subsets and a prognosis is made based on the expression of the genes, or their encoded proteins, in the geneset for that subset in a breast cancer tumor sample taken from the individual. The patient's responsiveness to chemotherapy is then determined based on the patient's subset classification and the prognosis. Various methods are known in the art that can be used for classifying an expression profile. Exemplary methods are described in Section 5.2.4, *infra*.

For example, a particular hypothetical condition has three relevant phenotypic characteristics, A, B and C. In this example, based on these characteristics, genesets informative for prognosis of four patient subsets A^+B^+ ; $A^+B^-C^+$; $A^-B^+C^-$; and A^- are identified by the method described above. Thus, an individual having the condition would first be classified according to phenotypes A-C into one of the four patient subsets. In one embodiment, therefore, the invention provides for the classification of an individual having a condition into one of a plurality of patient subsets, wherein a set of genes informative for prognosis for the subset has been identified. A sample is then taken from the individual, and the levels of expression of the prognostically-informative genes in the sample is analyzed. In one embodiment, the expression level of each gene can be compared to the expression level of the corresponding gene in a control or reference sample to determine a differential expression level. The expression profile comprising expression levels, e.g., differential expression levels, of the plurality of genes is then compared to a template profile. In various

embodiments, the template profile is a good prognosis template comprising the average expression of informative genes in samples taken from good prognosis individuals classifiable into that patient subset; or a poor prognosis template comprising the average expression of informative genes in samples taken from poor prognosis individuals classifiable into that patient subset; or a good prognosis profile comprising a set of mathematical values that represent gene expression levels of good prognosis individuals classifiable into that patient subset; etc.

In a specific embodiment, the phenotypic, genotypic and/or clinical classes for breast cancer are: ER⁻, *BRCA1* individuals; ER⁻, sporadic individuals; ER⁺, ER/AGE high individuals; ER⁺, ER/AGE low, LN⁺ individuals; and ER⁺, ER/AGE low, LN⁻ individuals. In this embodiment, an individual may be classified as ER⁺ or ER⁻. If the individual is ER⁻, the individual is additionally classified as having a *BRCA1*-type or sporadic tumor. ER⁻ individuals are thus classified as ER⁻, *BRCA1* or ER⁻, sporadic. Alternatively, if the individual is classified as ER⁺, the individual is additionally classified as having a high or low ratio of the log (ratio) of the level of expression of the gene encoding the estrogen receptor to the individual's age. Individuals having a low ratio are additionally classified as LN⁺ or LN⁻. ER⁺ individuals are thus classified as ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺, or ER⁺, ER/AGE low, LN⁻. Of course, the individual's ER status, tumor type, age and LN status may be identified in any order, as long as the individual is classified into one of these five subsets.

In one embodiment, a breast cancer patient is first classified into one of the following patient subsets: ER⁻, *BRCA1* individuals; ER⁻, sporadic individuals; ER⁺, ER/AGE high individuals; ER⁺, ER/AGE low, LN⁺ individuals; and ER⁺, ER/AGE low, LN⁻ individuals. The patient is then classified as having a good prognosis or a poor prognosis by comparing the patient's expression profile of a plurality of genes or their encoded proteins to a good and/or poor prognosis template profile of expression levels of said plurality of genes or their encoded proteins for the appropriate subset, and classifying said patient as having a good prognosis or poor prognosis if said patient expression profile has a high similarity to a good prognosis template or a poor prognosis template, respectively. In a preferred embodiment, the good prognosis template for a patient subset comprises the average level of expression of each of said plurality of genes informative for prognosis in tumor samples from individuals classified into said subset who have a good prognosis or good outcome, while the poor

prognosis template comprises the average level of expression of each of said plurality of genes informative for prognosis in tumor samples from individuals classified into said subset who have a poor prognosis or poor outcome. In another specific embodiment, said good or poor prognosis template is a set of mathematical values representing the average level of expression of genes informative for prognosis in tumor samples of individuals classifiable into said subset who have a good prognosis, or who have a poor prognosis, respectively.

It is evident that the different patient subsets described herein reflect different molecular mechanisms of the initiation of tumor formation and metastasis. Thus, the genesets listed in tables 1-5 are also useful for diagnosing a person as having a particular type of breast cancer in the first instance. Thus, the invention also provides a method of diagnosing an individual as having a particular subtype of breast cancer, comprising determining an expression profile of a plurality of the genes for which markers are listed in Tables 1-5 in a sample from said individual; and comparing said expression profile to a template profile, where said template is representative of the expression of said plurality of genes in a breast cancer sample of said subtype of cancer, and on the basis of said comparison, diagnosing the individual as having said subtype of breast cancer. In a specific embodiment, said subtype of cancer is selected from the group consisting of ER⁻, *BRCA1* type; ER⁻, sporadic type; ER⁺, ER/AGE high type; ER⁺, ER/AGE low, LN⁺ type; and ER/AGE low, LN⁻ type. In another specific embodiment, said template comprises the average levels of expression of a plurality of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5. In another specific example, said comparing comprises determining the similarity of the expression profile of the genes for which markers are listed in each of Tables 1-5 in said sample taken from said individual to a template profile comprising levels of expression of the same genes for each of Tables 1-5, and determining whether the pattern of expression of said genes in said sample is most similar to the pattern of expression of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5.

In another embodiment, a breast cancer patient is classified as having a good prognosis or a poor prognosis by a method comprising: (a) classifying said patient as ER⁻, *BRCA1*; ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining an expression profile of a first plurality of genes in a cell sample taken from the patient, said first plurality of genes comprising at least two of the genes

corresponding to the markers Table 1 if said patient is classified as ER⁻, *BRCA1*; Table 2 if said patient is classified as ER⁻, sporadic; Table 3 if said patient is classified as ER⁺, ER/AGE high; Table 4 if said patient is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said patient is classified as ER⁺, ER/AGE low, LN⁻, wherein said patient is "ER/AGE high" if the ER level of the patient exceeds a predetermined value, and "ER/AGE low" if the ER level of the patient does not exceed said predetermined value. In a specific embodiment of this method, said predetermined value of ER is calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said patient. In another specific embodiment, said patient is ER⁻, *BRCA1*, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 1. In another specific embodiment, said patient is ER⁻, sporadic, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 2. In another specific embodiment, said patient is ER⁺, ER/AGE high, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 3. In another specific embodiment, said patient is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 4. In another specific embodiment, said patient is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 5. In another specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

Where information is available regarding the LN status of a breast cancer patient, the patient may be identified as having a "very good prognosis," an "intermediate prognosis," or a poor prognosis, which enables the refinement of treatment. In one embodiment, the invention provides a method of assigning a therapeutic regimen to a breast cancer patient, comprising: (a) classifying said patient as having a "poor prognosis," "intermediate prognosis," or "very good prognosis" on the basis of the levels of expression of at least five genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5; and (b) assigning said patient a therapeutic regimen, said therapeutic regimen (i) comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or (ii) comprising chemotherapy if said patient has any other combination of lymph node status and expression profile.

In another embodiment, a breast cancer patient is assigned a prognosis by a method comprising (a) determining the breast cancer patient's age, ER status, LN status and tumor type; (b) classifying said patient as ER⁻, sporadic; ER⁻, *BRCAl*; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (c) determining an expression profile
5 comprising at least five genes in a cell sample taken from said breast cancer patient wherein markers for said at least five genes are listed in Table 1 if said patient is classified as ER⁻, sporadic; Table 2 if said patient is classified as ER⁻, *BRCAl*; Table 3 if said patient is classified as ER⁺, ER/AGE high; Table 4 if said patient is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said patient is classified as ER⁺, ER/AGE high, LN⁻; (d) determining the
10 similarity of the expression profile of said at least five genes to a template profile comprising levels of expression of said at least five genes to obtain a patient similarity value; (e) comparing said patient similarity value to selected first and second threshold values of similarity, respectively, wherein said second similarity threshold indicates greater similarity to said template expression profile than does said first similarity threshold; and (f) classifying
15 said breast cancer patient as having a first prognosis if said patient similarity value exceeds said second threshold similarity values, a second prognosis if said patient similarity value exceeds said first threshold similarity value but does not exceed said second threshold similarity value, and a third prognosis if said patient similarity value does not exceed said first threshold similarity value. In a specific embodiment of the method, said first prognosis
20 is a "very good prognosis," said second prognosis is an "intermediate prognosis," and said third prognosis is a "poor prognosis," wherein said breast cancer patient is assigned a therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and
25 expression profile.

The invention also provides a method of assigning a therapeutic regimen to a breast cancer patient, comprising: (a) determining the lymph node status for said patient; (b) determining the expression of at least five genes for which markers are listed in Table 5 in a cell sample from said patient, thereby generating an expression profile; (c) classifying said
30 patient as having a "poor prognosis," "intermediate prognosis," or "very good prognosis" on the basis of said expression profile; and (d) assigning said patient a therapeutic regimen, said therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or

comprising chemotherapy if said patient has any other combination of lymph node status and classification. In a specific embodiment of this method, said therapeutic regimen assigned to lymph node negative patients classified as having an "intermediate prognosis" additionally comprises adjuvant hormonal therapy. In another specific embodiment of this method, said

5 classifying step (c) is carried out by a method comprising: (a) rank ordering in descending order a plurality of breast cancer tumor samples that compose a pool of breast cancer tumor samples by the degree of similarity between the expression profile of said at least five genes in each of said tumor samples and the expression profile of said at least five genes across all remaining tumor samples that compose said pool, said degree of similarity being expressed as

10 a similarity value; (b) determining an acceptable number of false negatives in said classifying step, wherein a false negative is a breast cancer patient for whom the expression levels of said at least five genes in said cell sample predicts that said breast cancer patient will have no distant metastases within the first five years after initial diagnosis, but who has had a distant metastasis within the first five years after initial diagnosis; (c) determining a similarity value

15 above which in said rank ordered list said acceptable number of tumor samples or fewer are false negatives; (d) selecting said similarity value determined in step (c) as a first threshold similarity value; (e) selecting a second similarity value, greater than said first similarity value, as a second threshold similarity value; and (f) determining the similarity between the expression profile of said at least five genes in a breast cancer tumor sample from the breast

20 cancer patient and the expression profile of said respective at least five genes in said pool, to obtain a patient similarity value, wherein if said patient similarity value equals or exceeds said second threshold similarity value, said patient is classified as having a "very good prognosis"; if said patient similarity value equals or exceeds said first threshold similarity value, but is less than said second threshold similarity value, said patient is classified as

25 having an "intermediate prognosis"; and if said patient similarity value is less than said first threshold similarity value, said patient is classified as having a "poor prognosis." Another specific embodiment of this method comprises determining the estrogen receptor (ER) status of said patient, wherein if said patient is ER positive and lymph node negative, said therapeutic regimen assigned to said patient additionally comprises adjuvant hormonal

30 therapy.

The invention provides a method of predicting the responsiveness of a breast cancer patient to chemotherapy based on the patient's cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from the patient, the

patient's estrogen receptor level ("ER level"), and the patient's estrogen receptor level relative to the patient's age ("ER/AGE"). In the method of the invention, a patient is classified into an appropriate chemotherapy responsiveness group as compared to patients in the general population of breast cancer patients. The inventors have discovered that a patient

5 exhibits enhanced response to chemotherapy, e.g., to treatment by a combination of cyclophosphamide, methotrexate, and 5-fluorouracil (the "CMF combination"), as compared to patients in the general population of breast cancer patients if the patient's cellular constituent profile indicates poor prognosis, and the patient's ER level is ER⁺, and the patient's estrogen receptor level relative to the patient's age ER/AGE is low, whereas the

10 patient exhibits reduced response to chemotherapy as compared to patients in the general population if (i) the patient's ER level is ER⁻, or (ii) the patient's cellular constituent profile indicates poor prognosis, and the patient's ER level is ER⁺, and the patient's estrogen receptor level relative to the patient's age ER/AGE is high. Here, ER⁺ designates a high ER level and ER⁻ designates a low ER level. In one embodiment, the ER levels are measured by

15 immunohistochemical staining with tissue samples, and those patients whose tissue samples show greater than 10% of nuclei showing staining is deemed ER⁺ and less than 10% of nuclei showing staining is deemed ER⁻. In another embodiment, ER⁺ and ER⁻ patients are separated by the transcript or mRNA level of a gene encoding the estrogen receptor measured using a microarray, with ER⁺ patients having $\log_{10}(\text{ratio}) > -0.65$, and ER⁻ patients having

20 $\log_{10}(\text{ratio}) = \text{or } < -0.65$, where the ratio is the ratio of the estrogen receptor mRNA level of the patient sample and a control sample, e.g., a pool of cellular constituents from a plurality of different breast tumor samples. Preferably, the plurality of samples comprises at least 50, 100, 200, or 300 different samples. In addition, the inventors have also found that if the patient's cellular constituent profile indicates a good prognosis, then the patient does not need

25 chemotherapy. The method of the invention is particularly useful for predicting the responsiveness of a breast cancer patient under the age of 55 to chemotherapy, e.g., treatment by the CMF combination.

In one embodiment, the patient's cellular constituent profile is evaluated to determine whether the profile indicates good prognosis or poor prognosis. In a preferred embodiment,

30 the patient's prognosis is evaluated by comparing the cellular constituent profile to a predetermined cellular constituent template profile corresponding to a certain prognosis level, e.g., a good prognosis template comprising measurements of the plurality of cellular constituents which are representative of levels of the cellular constituents in a plurality of

good outcome patients or a poor prognosis template comprising measurements of the plurality of cellular constituents which are representative of levels of the cellular constituents in a plurality of poor outcome patients. In a preferred embodiment, the good prognosis template comprises average levels of the respective cellular constituents in a plurality of good outcome patients. In a preferred embodiment, the poor prognosis template comprises average levels of the respective cellular constituents in a plurality of poor outcome patients. In one embodiment, the average level of each cellular constituent in the good prognosis or poor prognosis template is a simple average. In another embodiment, the average level of each cellular constituent in the good prognosis or poor prognosis template is an error-weighted average. Herein a good outcome patient is a patient who has non-occurrence of metastases within a period of time after initial diagnosis, e.g., a period of 1, 2, 3, 4, 5 or 10 years, and a poor outcome patient is a patient who has occurrence of metastases within a period of time after initial diagnosis, e.g., a period of 1, 2, 3, 4, 5 or 10 years. In a preferred embodiment, both periods are 10 years. Table 8 shows exemplary good (column C₁) and poor (column C₂) template profiles for the 70-gene marker set.

The degree of similarity of the patient's cellular constituent profile to a template representing good or poor prognosis can be used to indicate whether the patient has good or poor prognosis. In a preferred embodiment, a patient is classified as having a good prognosis profile if the patient's cellular constituent profile has a high similarity to a good prognosis template and/or has a low similarity to a poor prognosis template. In another embodiment, a patient is classified as having a poor prognosis profile if the patient's cellular constituent profile has a low similarity to a good prognosis template and/or has a high similarity to a poor prognosis template. In embodiments for predicting the responsiveness of a breast cancer patient under the age of 55, the patients in the good and poor outcome patient populations used to generate the templates are preferably also under the age of 55 at the time of diagnosis of breast cancer.

The degree of similarity between a patient's cellular constituent profile and a template profile can be determined using any method known in the art. In one embodiment, the similarity is represented by a correlation coefficient between the patient's profile and the template. In one embodiment, a correlation coefficient above a correlation threshold indicates high similarity, whereas a correlation coefficient below the threshold indicates low similarity. In preferred embodiments, the correlation threshold is set as 0.3, 0.4, 0.5 or 0.6.

In another embodiment, similarity between a patient's profile and a template is represented by a distance between the patient's profile and the template. In one embodiment, a distance below a given value indicates high similarity, whereas a distance equal to or greater than the given value indicates low similarity.

- 5 As an illustration, in one embodiment, a template for a good prognosis is defined as \bar{c}_1 (e.g., a profile consisting of the values listed in the good prognosis C_1 column of Table 8) and/or a template for poor prognosis is defined as \bar{c}_2 (e.g., a profile consisting of the values listed in the good prognosis C_2 column of Table 8). Either one or both of the two classifier parameters (P_1 and P_2) can then be used to measure degrees of similarities between a
- 10 patient's profile and the templates: P_1 measures the similarity between the patient's profile \bar{y} and the good prognosis template \bar{c}_1 , and P_2 measures the similarity between \bar{y} and the poor prognosis template \bar{c}_2 . In embodiments which employ correlation coefficients, the correlation coefficients P_i can be calculated as:

$$P_i = (\bar{c}_i \cdot \bar{y}) / (\|\bar{c}_i\| \cdot \|\bar{y}\|) \quad (4)$$

- 15 where $i = 1$ and 2 .

Thus, in one embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a selected correlation threshold or if P_2 is equal to or less than a selected correlation threshold. In another embodiment, \bar{y} is classified as a poor prognosis profile if P_1 is less than a selected correlation threshold or if P_2 is above a selected correlation threshold. In still another

20 embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a first selected correlation threshold and \bar{y} is classified as a poor prognosis profile if P_2 is greater than a second selected correlation threshold.

Other methods known in the art, e.g., the methods described in Section 5.2.4, *infra*, can also be used for classifying an expression profile according to prognosis.

- 25 In a preferred embodiment, the cellular constituent profile is an expression profile comprising measurements of a plurality of transcripts (measured e.g., by measuring mRNAs or cDNA) in a sample derived from a patient. In this embodiment, the good prognosis template can be a good prognosis expression template comprising measurements of the

plurality of transcripts which are representative of expression levels of the transcripts in a plurality of good outcome patients, and the poor prognosis template can be a poor prognosis expression template comprising measurements of the plurality of transcripts which are representative of expression levels of the transcripts in a plurality of poor outcome patients.

- 5 In a preferred embodiment, measurement of each transcript in the good or poor prognosis expression template is an average of expression levels of the transcript in the plurality of good or poor outcome patients, respectively. In one embodiment, each measurement is a mean subtracted log(intensity) (see, PCT publication WO 2004/065545, published August 5, 2004, which is incorporated herein by reference in its entirety). In one embodiment, the
- 10 plurality of transcripts consists of transcripts corresponding to at least a portion of the set of genes listed in Table 8. Preferably, the plurality of transcripts consists of transcripts corresponding to at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the set of genes listed in Table 8, or at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, or 65 of the genes listed in Table 8. In a preferred embodiment, the plurality of transcripts consists of
- 15 transcripts corresponding to all the genes listed in Table 8.

- In another embodiment, the expression profile is a differential expression profile comprising differential measurements of said plurality of transcripts in a sample derived from said patient versus measurements of said plurality of transcripts in a control sample. The differential measurements can be x_{dev} , log(ratio), error-weighted log(ratio), or a mean
- 20 subtracted log(intensity) (see, e.g., PCT publication WO 00/39339, published on July 6, 2000; PCT publication WO 2004/065545, published August 5, 2004, each of which is incorporated herein by reference in its entirety).

- In another embodiment, the cellular constituent profile comprises measurements of a plurality of protein species in a sample derived from the patient. In this embodiment, a good
- 25 prognosis template comprises measurements of the plurality of protein species which are representative of levels of the protein species in a plurality of good outcome patients, and the poor prognosis template comprises measurements of the plurality of protein species which are representative of levels of the protein species in the plurality of poor outcome patients.

- The patient's cellular constituent profile can be determined by any method known in
- 30 the art, e.g., as described in Section 5.3.

The methods of the invention preferably use a control or reference sample, which can be any suitable sample against which changes in cellular constituents can be determined. In one embodiment, the control or reference sample is generated by pooling together the plurality of cellular constituents, e.g., a plurality of transcripts or cDNAs, or a plurality of protein species, from a plurality of breast cancer patients. Alternatively, the control or reference sample can be generated by pooling together purified or synthesized cellular constituents, e.g., a plurality of purified or synthesized transcripts or cDNAs, a plurality of purified or synthesized protein species. In one embodiment, synthetic RNAs for each transcripts or cDNAs are pooled to form the control or reference sample. Preferably, the abundances of synthetic RNAs are approximately the abundances of the corresponding transcripts in a real tumor pool. The differential expression of marker genes for each individual patient sample is measured against this control sample. In one embodiment, 60-mer oligonucleotides corresponding to the probe sequences on a microarray used to assay the expression levels of the diagnostic/prognostic transcripts are synthesized and cloned into pBluescript SK- vector (Statagene, La Jolla, CA), adjacent to the T7 promoter sequence. Individual clones are isolated, and the sequences of their inserts are verified by DNA sequencing. To generate synthetic RNAs, clones are linearized with *EcoRI* and a T7 *in vitro* transcription (IVT) reaction is performed by MegaScript kit (Ambion, Austin, TX), followed by DNase treatment of the product. Synthetic RNAs are purified on RNeasy columns (Qiagen, Valencia, CA). These synthetic RNAs are transcribed, amplified, labeled, and mixed together to make the reference pool. The abundance of those synthetic RNAs are chosen to approximate the abundances of the transcripts of the corresponding marker genes in the real tumor pool.

In one embodiment, the ER level is determined by measuring an expression level of a gene encoding the estrogen receptor in the patient relative to that in a control sample. In a preferred embodiment, the gene encoding the estrogen receptor is the estrogen receptor α gene. In another embodiment, the ER level is measured by an oligonucleotide probe that detects the transcript of the gene having accession number NM_000125 (SEQ ID NO: 425). The control sample is a pool of breast cancer cells from different patients. In this embodiment, the ER level is classified as ER⁺ if the log(ratio) > -0.65, and is classified as ER⁻ otherwise (see, e.g., van't Veer et al., 2002, Nature 415:530). In another embodiment, the ER level is determined by immunohistochemical staining measurement of the level of the estrogen receptors in a tissue sample of the patient. In one embodiment, the ER level is

classified as ER⁺ if immunohistochemical staining shows greater than 10% of nuclei showing staining and the ER level is classified as ER⁺, if less than 10% nuclei of showing staining.

The patient's estrogen receptor level relative to the patient's age can be measured using an appropriate metric designated as ER/AGE. If the individual's ER level is high relative to the individual's age, e.g., if the expression level of an estrogen receptor gene of the individual is higher than an age-dependent threshold value, the individual is classified as ER/AGE high. In one embodiment, the log(ratio) of ER expression of an individual of a particular age is compared with the threshold value for the individual's age, and an ER expression level equal to or greater than the threshold value is deemed high. In one embodiment, the patient's estrogen receptor level as measured by log10(ratio) is compared to a parameter defined as $c \cdot (\text{AGE} - d)$, where c is a coefficient, AGE is the age of said patient, and d is an age threshold. The patient is classified as having high ER/AGE if the patient's ER level is greater than $c \cdot (\text{AGE} - d)$, and the patient is classified as having low ER/AGE if the patient's ER level is equal to or less than $c \cdot (\text{AGE} - d)$. The parameters c and d depend on the ER level and AGE used. They can be determined by fitting patients' ER level-age distribution to a bimodal distribution of two subgroups each having a different ER level-age dependence. In a preferred embodiment, $c = 0.1$ and $d = 42.5$.

Chemotherapy can be performed using any one or a combination of the anti-cancer drugs known in the art, including but are not limited to any topoisomerase inhibitor, DNA binding agent, anti-metabolite, ionizing radiation, or a combination of two or more of such known DNA damaging agents.

A topoisomerase inhibitor that can be used in conjunction with the invention can be a topoisomerase I (Topo I) inhibitor, a topoisomerase II (Topo II) inhibitor, or a dual topoisomerase I and II inhibitor. A topo I inhibitor can be from any of the following classes of compounds: camptothecin analogue (e.g., karenitecin, aminocamptothecin, lurtotecan, topotecan, irinotecan, BAY 56-3722, rubitecan, GI14721, exatecan mesylate), rebeccamycin analogue, PNU 166148, rebeccamycin, TAS-103, camptothecin (e.g., camptothecin polyglutamate, camptothecin sodium), intoplicine, ecteinascidin 743, J-107088, pibenzimol. Examples of preferred topo I inhibitors include but are not limited to camptothecin, topotecan (hycaptamine), irinotecan (irinotecan hydrochloride), belotecan, or an analogue or derivative thereof.

A topo II inhibitor that can be used in conjunction with the invention can be from any of the following classes of compounds: anthracycline antibiotics (e.g., carubicin, pirarubicin, daunorubicin citrate liposomal, daunomycin, 4-iodo-4'-doxydoxorubicin, doxorubicin, n,n-dibenzyl daunomycin, morpholinodoxorubicin, aclacinomycin antibiotics, duborimycin, menogaril, nogalamycin, zorubicin, epirubicin, marcellomycin, detorubicin, annamycin, 7-cyanoquinocarcinol, deoxydoxorubicin, idarubicin, GPX-100, MEN-10755, valrubicin, KRN5500), epipodophyllotoxin compound (e.g., podophyllin, teniposide, etoposide, GL331, 2-ethylhydrazide), anthraquinone compound (e.g., ametantrone, bisantrene, mitoxantrone, anthraquinone), ciprofloxacin, acridine carboxamide, amonafide, anthrapyrazole antibiotics (e.g., teloxantrone, sedoxantrone trihydrochloride, piroxantrone, anthrapyrazole, losoxantrone), TAS-103, fostriecin, razoxane, XK469R, XK469, chloroquinoxaline sulfonamide, merbarone, intoplicine, elsamitrucin, CI-921, pyrazoloacridine, elliptinium, amsacrine. Examples of preferred topo II inhibitors include but are not limited to doxorubicin (Adriamycin), etoposide phosphate (etopofos), teniposide, sobuzoxane, or an analogue or derivative thereof.

DNA binding agents that can be used in conjunction with the invention include but are not limited to DNA groove binding agent, e.g., DNA minor groove binding agent; DNA crosslinking agent; intercalating agent; and DNA adduct forming agent. A DNA minor groove binding agent can be an anthracycline antibiotic, mitomycin antibiotic (e.g., porfiromycin, KW-2149, mitomycin B, mitomycin A, mitomycin C), chromomycin A3, carzelesin, actinomycin antibiotic (e.g., cactinomycin, dactinomycin, actinomycin F1), brostallicin, echinomycin, bizelesin, duocarmycin antibiotic (e.g., KW 2189), adozelesin, olivomycin antibiotic, plicamycin, zinostatin, distamycin, MS-247, ecteinascidin 743, amsacrine, anthramycin, and pibenzimol, or an analogue or derivative thereof.

DNA crosslinking agents include but are not limited to antineoplastic alkylating agent, methoxsalen, mitomycin antibiotic, psoralen. An antineoplastic alkylating agent can be a nitrosourea compound (e.g., cystemustine, taumustine, semustine, PCNU, streptozocin, SarCNU, CGP-6809, carmustine, fotemustine, methyl nitrosourea, nimustine, ranimustine, ethyl nitrosourea, lomustine, chlorozotocin), mustard agent (e.g., nitrogen mustard compound, such as spiromustine, trofosfamide, chlorambucil, estramustine, 2,2,2-trichlorotriethylamine, prednimustine, novembichin, phenamet, glufosfamide, peptichemio, ifosfamide, defosfamide, nitrogen mustard, phenesterin, mannomustine, cyclophosphamide,

- melphalan, perfosfamide, mechlorethamine oxide hydrochloride, uracil mustard, bestrabucil, DHEA mustard, tallimustine, mafosfamide, aniline mustard, chlornaphazine; sulfur mustard compound, such as bischloroethylsulfide; mustard prodrug, such as TLK286 and ZD2767), ethylenimine compound (e.g., mitomycin antibiotic, ethylenimine, uredepa, thiotepa,
- 5 diaziquone, hexamethylene bisacetamide, pentamethylmelamine, altretamine, carzinophilin, triaziquone, meturedepa, benzodepa, carboquone), alkylsulfonate compound (e.g., dimethylbusulfan, Yoshi-864, improsulfan, piposulfan, treosulfan, busulfan, hepsulfam), epoxide compound (e.g., anaxirone, mitolactol, dianhydrogalactitol, teroxirone), miscellaneous alkylating agent (e.g., ipomeanol, carzelesin, methylene dimethane sulfonate,
- 10 mitobronitol, bizelesin, adozelesin, piperazinedione, VNP40101M, asaley, 6-hydroxymethylacetylfulvene, EO9, etoglucid, ecteinascidin 743, pipobroman), platinum compound (e.g., ZD0473, liposomal-cisplatin analogue, satraplatin, BBR 3464, spiroplatin, ormaplatin, cisplatin, oxaliplatin, carboplatin, lobaplatin, zeniplatin, iproplatin), triazene compound (e.g., imidazole mustard, CB10-277, mitozolomide, temozolomide, procarbazine,
- 15 dacarbazine), picoline compound (e.g., penclomedine), or an analogue or derivative thereof. Examples of preferred alkylating agents include but are not limited to cisplatin, dibromodulcitol, fotemustine, ifosfamide (ifosfamid), ranimustine (ranomustine), nedaplatin (latoplatin), bendamustine (bendamustine hydrochloride), eptaplatin, temozolomide (methazolastone), carboplatin, altretamine (hexamethylmelamine), prednimustine, oxaliplatin
- 20 (oxalaplatinum), carmustine, thiotepa, leusulfon (busulfan), lobaplatin, cyclophosphamide, bisulfan, melphalan, and chlorambucil, or analogues or derivatives thereof.

- Intercalating agents can be an anthraquinone compound, bleomycin antibiotic, rebeccamycin analogue, acridine, acridine carboxamide, amonafide, rebeccamycin, anthrapyrazole antibiotic, echinomycin, psoralen, LU 79553, BW A773U, crisnatol mesylate,
- 25 benzo(a)pyrene-7,8-diol-9,10-epoxide, acodazole, elliptinium, pixantrone, or an analogue or derivative thereof.

- DNA adduct forming agents include but are not limited to enediyne antitumor antibiotic (e.g., dynemicin A, esperamicin A1, zinostatin, dynemicin, calicheamicin gamma 1I), platinum compound, carmustine, tamoxifen (e.g., 4-hydroxy-tamoxifen), psoralen,
- 30 pyrazine diazohydroxide, benzo(a)pyrene-7,8-diol-9,10-epoxide, or an analogue or derivative thereof.

Anti-metabolites include but are not limited to cytosine, arabinoside, floxuridine, fluorouracil, mercaptopurine, Gemcitabine, and methotrexate (MTX).

- In a specific embodiment, a KSP inhibitor (1S)-1-[[[(2S)-4-(2,5-difluorophenyl)-2-phenyl-2,5-dihydro-1H-pyrrol-1-yl]carbonyl]-2-methylpropylamine (see, PCT application 5 PCT/US03/18482, filed June 12, 2003, which is incorporated herein by reference in its entirety), is used alone or in combination with other anti-cancer drugs for chemotherapy.

In another specific embodiment, the chemotherapy is carried out using the CMF combination consisting of cyclophosphamide, methotrexate, and 5-fluorouracil.

5.2.3. IMPROVING SENSITIVITY TO EXPRESSION LEVEL DIFFERENCES

- 10 In using the markers disclosed herein, and, indeed, using any sets of markers, e.g., to compare profiles or to differentiate an individual having one phenotype from another individual having a second phenotype, one can compare the profile comprising absolute expression levels of the markers in a sample to a template; for example, a template comprising the average levels of expression of the markers in a plurality of individuals. To 15 increase the sensitivity of the comparison, however, the expression level values are preferably transformed in a number of ways. Also, to differentiate an individual having one phenotype from another individual having a second phenotype using any sets of markers, one can compare the absolute expression of each of the markers in a sample to a control; for example, the control can be the average level of expression of each of the markers, respectively, in a 20 pool of individuals.

- For example, the expression level of each of the markers can be normalized by the average expression level of all markers the expression level of which is determined, or by the average expression level of a set of control genes. Thus, in one embodiment, the markers are represented by probes on a microarray, and the expression level of each of the markers is 25 normalized by the mean or median expression level across all of the genes represented on the microarray, including any non-marker genes. In a specific embodiment, the normalization is carried out by dividing the median or mean level of expression of all of the genes on the microarray. In another embodiment, the expression levels of the markers is normalized by the mean or median level of expression of a set of control markers. In a specific 30 embodiment, the control markers comprise a set of housekeeping genes. In another specific

embodiment, the normalization is accomplished by dividing by the median or mean expression level of the control genes.

The sensitivity of a marker-based assay will also be increased if the expression levels of individual markers are compared to the expression of the same markers in a pool of samples. Preferably, the comparison is to the mean or median expression level of each the marker genes in the pool of samples. Such a comparison may be accomplished, for example, by dividing by the mean or median expression level of the pool for each of the markers from the expression level each of the markers in the sample. This has the effect of accentuating the relative differences in expression between markers in the sample and markers in the pool as a whole, making comparisons more sensitive and more likely to produce meaningful results that the use of absolute expression levels alone. The expression level data may be transformed in any convenient way; preferably, the expression level data for all is log transformed before means or medians are taken.

In performing comparisons to a pool, two approaches may be used. First, the expression levels of the markers in the sample may be compared to the expression level of those markers in the pool, where nucleic acid derived from the sample and nucleic acid derived from the pool are hybridized during the course of a single experiment. Such an approach requires that new pool nucleic acid be generated for each comparison or limited numbers of comparisons, and is therefore limited by the amount of nucleic acid available. Alternatively, and preferably, the expression levels in a pool, whether normalized and/or transformed or not, are stored on a computer, or on computer-readable media, to be used in comparisons to the individual expression level data from the sample (i.e., single-channel data).

The current invention also provides the following method of classifying a first cell or organism as having one of at least two different phenotypes, where the different phenotypes comprise a first phenotype and a second phenotype. The level of expression of each of a plurality of markers in a first sample from the first cell or organism is compared to the level of expression of each of said markers, respectively, in a pooled sample from a plurality of cells or organisms, the plurality of cells or organisms comprising different cells or organisms exhibiting said at least two different phenotypes, respectively, to produce a first compared value. The first compared value is then compared to a second compared value, wherein said second compared value is the product of a method comprising comparing the level of

expression of each of said markers in a sample from a cell or organism characterized as having said first phenotype to the level of expression of each of said markers, respectively, in the pooled sample. The first compared value is then compared to a third compared value, wherein said third compared value is the product of a method comprising comparing the level of expression of each of the markers in a sample from a cell or organism characterized as having the second phenotype to the level of expression of each of the markers, respectively, in the pooled sample. In specific embodiments, the marker can be a gene, a protein encoded by the gene, etc. Optionally, the first compared value can be compared to additional compared values, respectively, where each additional compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having a phenotype different from said first and second phenotypes but included among the at least two different phenotypes, to the level of expression of each of said genes, respectively, in said pooled sample. Finally, a determination is made as to which of said second, third, and, if present, one or more additional compared values, said first compared value is most similar, wherein the first cell or organism is determined to have the phenotype of the cell or organism used to produce said compared value most similar to said first compared value.

The sensitivity of a marker-based assay will also be increased if the expression levels of individual markers are compared to the expression of the same markers in a control sample, e.g., a sample comprises a pool of samples, to generate a differential expression profile. Such a comparison may be accomplished, for example, by determining a ratio between expression level of each marker in the sample and the expression level of the corresponding marker in the control sample. This has the effect of accentuating the relative differences in expression between markers in the sample and markers in the control as a whole, making subsequent comparisons to a template more sensitive and more likely to produce meaningful results than the use of absolute expression levels alone. The comparison may be performed in any convenient way, e.g., by taking difference, ratio, or log(ratio).

In performing comparisons to a control sample, two approaches may be used. First, the expression levels of the markers in the sample may be compared to the expression level of those markers in the control sample, e.g., where nucleic acid derived from the sample and nucleic acid derived from the control are hybridized during the course of a single experiment. Such an approach requires that new control sample of nucleic acid be generated for each

comparison or limited numbers of comparisons, and is therefore limited by the amount of nucleic acid available. Alternatively, the expression levels in a control sample, whether normalized and/or transformed or not, are stored on a computer, or on computer-readable media, to be used in comparisons to the individual expression level data from the sample (i.e., single-channel data).

The methods of the invention preferably use a control or reference sample, which can be any suitable sample against which changes in cellular constituents can be determined. In one embodiment, the control or reference sample is generated by pooling together the plurality of cellular constituents, e.g., a plurality of transcripts or cDNAs, or a plurality of protein species, from a plurality of breast cancer patients. Alternatively, the control or reference sample can be generated by pooling together purified or synthesized cellular constituents, e.g., a plurality of purified or synthesized transcripts or cDNAs, a plurality of purified or synthesized protein species. In one embodiment, synthetic RNAs for each transcripts or cDNAs are pooled to form the control or reference sample. Preferably, the abundances of synthetic RNAs are approximately the abundances of the corresponding transcripts in a real tumor pool. The differential expression of marker genes for each individual patient sample is measured against this control sample. In one embodiment, 60-mer oligonucleotides corresponding to the probe sequences on a microarray used to assay the expression levels of the diagnostic/prognostic transcripts are synthesized and cloned into pBluescript SK- vector (Statagene, La Jolla, CA), adjacent to the T7 promotor sequence. Individual clones are isolated, and the sequences of their inserts are verified by DNA sequencing. To generate synthetic RNAs, clones are linearized with *EcoRI* and a T7 in vitro transcription (IVT) reaction is performed by MegaScript kit (Ambion, Austin, TX), followed by DNase treatment of the product. Synthetic RNAs are purified on RNeasy columns (Qiagen, Valencia, CA). These synthetic RNAs are transcribed, amplified, labeled, and mixed together to make the reference pool. The abundance of those synthetic RNAs are chosen to approximate the abundances of the transcripts of the corresponding marker genes in the real tumor pool.

The current invention provides the following method of classifying a first cell or organism as having one of at least two different phenotypes, where the different phenotypes comprise a first phenotype and a second phenotype. The level of expression of each of a plurality of markers in a first sample from the first cell or organism is compared to the level

of expression of each of said markers, respectively, in a pooled sample from a plurality of cells or organisms, the plurality of cells or organisms comprising different cells or organisms exhibiting said at least two different phenotypes, respectively, to produce a first compared value so that a first differential profile comprising a plurality of first compared values for said plurality of markers is generated. The first differential profile is then compared to a second differential profile comprising a plurality of second compared values, wherein each said second compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having said first phenotype to the level of expression of each of said markers, respectively, in the pooled sample. The first differential profile is then compared to a third differential profile comprising a plurality of third compared values, wherein each said third compared value is the product of a method comprising comparing the level of expression of each of the markers in a sample from a cell or organism characterized as having the second phenotype to the level of expression of each of the markers, respectively, in the pooled sample. In specific embodiments, each marker can be a gene, a protein encoded by the gene, etc. Optionally, the first differential profile can be compared to additional expression profiles each of which comprising additional compared values, respectively, where each additional compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having a phenotype different from said first and second phenotypes but included among the at least two different phenotypes, to the level of expression of each of said genes, respectively, in said pooled sample. Finally, a determination is made as to which of said second, third, and, if present, one or more additional differential profiles, said first differential profile is most similar, wherein the first cell or organism is determined to have the phenotype of the cell or organism used to produce said differential profile most similar to said first differential profile.

In a specific embodiment of this method, the compared values are each ratios of the levels of expression of each of said genes. In another specific embodiment, each of the levels of expression of each of the genes in the pooled sample is normalized prior to any of the comparing steps. In a more specific embodiment, the normalization of the levels of expression is carried out by dividing by the median or mean level of the expression of each of the genes or dividing by the mean or median level of expression of one or more housekeeping genes in the pooled sample from said cell or organism. In another specific embodiment, the normalized levels of expression are subjected to a log transform, and the comparing steps

comprise subtracting the log transform from the log of the levels of expression of each of the genes in the sample. In another specific embodiment, the two or more different phenotypes are different stages of a disease or disorder. In still another specific embodiment, the two or more different phenotypes are different prognoses of a disease or disorder. In yet another specific embodiment, the levels of expression of each of the genes, respectively, in the pooled sample or said levels of expression of each of said genes in a sample from the cell or organism characterized as having the first phenotype, second phenotype, or said phenotype different from said first and second phenotypes, respectively, are stored on a computer or on a computer-readable medium.

In another specific embodiment, the two phenotypes are good prognosis and poor prognosis. In a more specific embodiment, the two phenotypes are good prognosis and poor prognosis for an individual that is identified as having ER⁻, *BRCA1* status, ER⁻, sporadic status, ER⁺, ER/AGE high status, ER⁺, ER/AGE low, LN⁺ status, or ER⁺, ER/AGE low, LN⁺ status.

In another specific embodiment, the comparison is made between the expression profile of the genes in the sample and the expression profile of the same genes in a pool representing only one of two or more phenotypes. In the context of prognosis-correlated genes, for example, one can compare the expression levels of prognosis-related genes in a sample to the average levels of the expression of the same genes in a plurality of "good prognosis" samples (as opposed to a plurality of samples that include samples from patients having poor prognoses and good prognoses). Thus, in this method, a sample is classified as having a good prognosis if the expression profile of prognosis-correlated genes exceeds a chosen coefficient of correlation to the average "good prognosis" expression profile (e.g., the profile comprising average levels of expression of prognosis-correlated genes in samples from a plurality of patients having a "good prognosis"). Patients whose expression profiles correlate more poorly with the "good prognosis" expression profile (e.g., whose correlation coefficient fails to exceed the chosen coefficient) are classified as having a poor prognosis.

Where individuals are classified on the basis of phenotypic, genotypic, or clinical characteristics into patient subsets, the pool of samples may be a pool of samples for the phenotype that includes samples representing each of the patient subsets. Alternatively, the pool of samples may be a pool of samples for the phenotype representing only the specific patient subset. For example, where an individual is classified as ER⁺, sporadic, the pool of

samples to which the individual's sample is compared may be a pool of samples from ER+, sporadic individuals having a good prognosis only, or may be a pool of samples of individuals having a good prognosis, without regard to ER status or mutation type.

- The method can be applied to a plurality of patient subsets. For example, in a specific embodiment, the phenotype is good prognosis, and the individual is classified into one of the following patient subsets: ER⁻, *BRCA1* status, ER⁻, sporadic status, ER+, ER/AGE high status, ER+, ER/AGE low, LN+ status, or ER+, ER/AGE low, LN+ status. A set of markers informative for prognosis for the patient subset into which the individual is classified is then used to determine the likely prognosis for the individual. A sample is classified as coming from an individual having a good prognosis if the expression profile of prognosis-correlated genes for the particular subset into which the individual is classified exceeds a chosen coefficient of correlation to the average "good prognosis" expression profile (e.g., the levels of expression of prognosis-correlated genes in a plurality of samples from patients within the subclass having a "good prognosis"). Patients whose expression levels correlate more poorly with the "good prognosis" expression profile (e.g., whose correlation coefficient fails to exceed the chosen coefficient) are classified as having a poor prognosis.

- Of course, single-channel data may also be used without specific comparison to a mathematical sample pool. For example, a sample may be classified as having a first or a second phenotype, wherein the first and second phenotypes are related, by calculating the similarity between the expression profile of at least 5 markers in the sample, where the markers are correlated with the first or second phenotype, to a first phenotype template and a second phenotype template each comprising the expression levels of the same markers, by (a) labeling nucleic acids derived from a sample with a fluorophore to obtain a pool of fluorophore-labeled nucleic acids; (b) contacting said fluorophore-labeled nucleic acid with a microarray under conditions such that hybridization can occur, detecting at each of a plurality of discrete loci on the microarray a fluorescent emission signal from said fluorophore-labeled nucleic acid that is bound to said microarray under said conditions; and (c) determining the similarity of marker gene expression in the individual sample to the first and second templates, wherein if said expression is more similar to the first template, the sample is classified as having the first phenotype, and if said expression is more similar to the second template, the sample is classified as having the second phenotype.

In a specific embodiment of the above method, the first phenotype is a good prognosis of breast cancer, the sample is a sample from an individual that has been classified into a patient subset, and the first and second templates are templates for the phenotype for the particular patient subset. In a more specific embodiment, for example, the first phenotype is a good prognosis, the second phenotype is a poor prognosis, the patient is classified into an ER⁻, sporadic patient subset, an ER⁻, *BRCA1* subset, an ER⁺, ER/AGE high subset, an ER⁺, ER/AGE low, LN⁺ subset, or an ER⁺, ER/AGE low, LN⁺ subset, and said first and second templates are templates derived from the expression of the marker genes in individuals having a good prognosis and a poor prognosis, respectively, wherein said individuals are all of the patient subset into which said patient is classified.

5.2.4. METHODS FOR CLASSIFICATION OF EXPRESSION PROFILES

In preferred embodiments, the methods of the invention use a classifier for predicting prognosis in a patient. The classifier can be based on any appropriate pattern recognition method that receives an input comprising a marker profile and provides an output comprising data indicating which patient subset the patient belongs. The classifier can be trained with training data from a training population of breast cancer patients. Typically, the training data comprise for each of the patients in the training population a training marker profile comprising measurements of respective gene products of a plurality of genes in a suitable sample taken from the patient and prognosis information.

In preferred embodiments, the classifier can be based on a classification (pattern recognition) method described below, e.g., profile similarity (Section 5.2.4.1., *infra*); artificial neural network (Section 5.2.4.2., *infra*); support vector machine (SVM, Section 5.2.4.3., *infra*); logic regression (Section 5.2.4.4., *infra*), linear or quadratic discriminant analysis (Section 5.2.4.5., *infra*), decision trees (Section 5.2.4.6., *infra*), clustering (Section 5.2.4.7., *infra*), principal component analysis (Section 5.2.4.8., *infra*), nearest neighbor classifier analysis (Section 5.2.4.9., *infra*). Such classifiers can be trained with the training population using methods described in the relevant sections, *infra*.

The marker profile can be obtained by measuring the plurality of gene products in a cell sample from the patient using a method known in the art, e.g., a method described in Section 5.3., *infra*.

Various known statistical pattern recognition methods can be used in conjunction with the present invention. A classifier based on any of such methods can be constructed using the marker profiles and prognosis data of training patients. Such a classifier can then be used to evaluate the prognosis status of a patient based on the patient's marker profile. The methods
 5 can also be used to identify markers that discriminate between different prognosis status using a marker profile and prognosis data of training patients.

5.2.4.1. PROFILE MATCHING

A patient's prognosis can be evaluated by comparing a marker profile obtained in a suitable sample from the patient with a marker profile that is representative of a particular
 10 prognosis. Such a marker profile is also termed a "template profile" or a "template." The degree of similarity to such a template profile provides an evaluation of the patient's prognosis. If the degree of similarity of the patient marker profile and a template profile is above a predetermined threshold, the patient is assigned the prognosis represented by the template. For example, a patient's prognosis can be evaluated by comparing a marker profile
 15 of the patient to a predetermined template profile corresponding to a given prognosis, e.g., a good prognosis template comprising measurements of the plurality of markers which are representative of levels of the markers in a plurality of patients having a good prognosis.

In one embodiment, the similarity is represented by a correlation coefficient between the patient's profile and the template. In one embodiment, a correlation coefficient above a
 20 correlation threshold indicates a high similarity, whereas a correlation coefficient below the threshold indicates a low similarity.

In a specific embodiment, P_i measures the similarity between the patient's profile \bar{y} and a template profile comprising measurements of marker gene products representative of measurements of marker gene products in patients having a particular prognosis, e.g., the
 25 good prognosis template \bar{z}_1 or the poor prognosis template \bar{z}_2 . Such a coefficient, P_i , can be calculated using the following equation:

$$P_i = (\bar{z}_i \bullet \bar{y}) / (\|\bar{z}_i\| \cdot \|\bar{y}\|)$$

where i designates the i th template. Thus, in one embodiment, \bar{y} is classified as a good prognosis profile if P_i is greater than a selected correlation threshold. In another

embodiment, \bar{y} is classified as poor prognosis profile if P_2 is greater than a selected correlation threshold. In preferred embodiments, the correlation threshold is set as 0.3, 0.4, 0.5 or 0.6. In another embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than P_2 , whereas \bar{y} is classified as a poor prognosis profile if P_1 is less than P_2 .

- 5 In another embodiment, the correlation coefficient is a weighted dot product of the patient's profile \bar{y} and a template profile, in which measurements of each different marker is assigned a weight.

- In another embodiment, similarity between a patient's profile and a template is represented by a distance between the patient's profile and the template. In one embodiment, 10 a distance below a given value indicates high similarity, whereas a distance equal to or greater than the given value indicates low similarity.

In one embodiment, the Euclidian distance according to the formula

$$D_i = \|\bar{y} - \bar{z}_i\|$$

- is used, where D_i measures the distance between the patient's profile \bar{y} and a template 15 profile comprising measurements of marker gene products representative of measurements of marker gene products in patients having a particular prognosis, e.g., the good prognosis template \bar{z}_1 or the poor prognosis template \bar{z}_2 . In other embodiments, the Euclidian distance is squared to place progressively greater weight on cellular constituents that are further apart. In alternative embodiments, the distance measure D_i is the Manhattan distance provide by

$$20 \quad D_i = \sum_n |y(n) - z_i(n)|$$

where $y(n)$ and $z_i(n)$ are respectively measurements of the n th marker gene product in the patient's profile \bar{y} and a template profile.

In another embodiment, the distance is defined as $D_i = 1 - P_i$, where P_i is the correlation coefficient or normalized dot product as described above.

- 25 In still other embodiments, the distance measure may be the Chebychev distance, the power distance, and percent disagreement, all of which are well known in the art.

5.2.4.2. ARTIFICIAL NEURAL NETWORK

In some embodiments, a neural network is used. A neural network can be constructed for a selected set of molecular markers of the invention. A neural network is a two-stage regression or classification model. A neural network has a layered structure that includes a layer of input units (and the bias) connected by a layer of weights to a layer of output units. For regression, the layer of output units typically includes just one output unit. However, neural networks can handle multiple quantitative responses in a seamless fashion.

In multilayer neural networks, there are input units (input layer), hidden units (hidden layer), and output units (output layer). There is, furthermore, a single bias unit that is connected to each unit other than the input units. Neural networks are described in Duda *et al.*, 2001, *Pattern Classification*, Second Edition, John Wiley & Sons, Inc., New York; and Hastie *et al.*, 2001, *The Elements of Statistical Learning*, Springer-Verlag, New York.

The basic approach to the use of neural networks is to start with an untrained network, present a training pattern, e.g., marker profiles from training patients, to the input layer, and to pass signals through the net and determine the output, e.g., the prognosis in the training patients, at the output layer. These outputs are then compared to the target values; any difference corresponds to an error. This error or criterion function is some scalar function of the weights and is minimized when the network outputs match the desired outputs. Thus, the weights are adjusted to reduce this measure of error. For regression, this error can be sum-of-squared errors. For classification, this error can be either squared error or cross-entropy (deviation). See, e.g., Hastie *et al.*, 2001, *The Elements of Statistical Learning*, Springer-Verlag, New York.

Three commonly used training protocols are stochastic, batch, and on-line. In stochastic training, patterns are chosen randomly from the training set and the network weights are updated for each pattern presentation. Multilayer nonlinear networks trained by gradient descent methods such as stochastic back-propagation perform a maximum-likelihood estimation of the weight values in the model defined by the network topology. In batch training, all patterns are presented to the network before learning takes place. Typically, in batch training, several passes are made through the training data. In online training, each pattern is presented once and only once to the net.

In some embodiments, consideration is given to starting values for weights. If the weights are near zero, then the operative part of the sigmoid commonly used in the hidden layer of a neural network (see, e.g., Hastie *et al.*, 2001, *The Elements of Statistical Learning*, Springer-Verlag, New York) is roughly linear, and hence the neural network collapses into an approximately linear model. In some embodiments, starting values for weights are chosen to be random values near zero. Hence the model starts out nearly linear, and becomes nonlinear as the weights increase. Individual units localize to directions and introduce nonlinearities where needed. Use of exact zero weights leads to zero derivatives and perfect symmetry, and the algorithm never moves. Alternatively, starting with large weights often leads to poor solutions.

Since the scaling of inputs determines the effective scaling of weights in the bottom layer, it can have a large effect on the quality of the final solution. Thus, in some embodiments, at the outset all expression values are standardized to have mean zero and a standard deviation of one. This ensures all inputs are treated equally in the regularization process, and allows one to choose a meaningful range for the random starting weights. With standardization inputs, it is typical to take random uniform weights over the range $[-0.7, +0.7]$.

A recurrent problem in the use of networks having a hidden layer is the optimal number of hidden units to use in the network. The number of inputs and outputs of a network are determined by the problem to be solved. In the present invention, the number of inputs for a given neural network can be the number of molecular markers in the selected set of molecular markers of the invention. The number of output for the neural network will typically be just one. However, in some embodiment more than one output is used so that more than just two states can be defined by the network. If too many hidden units are used in a neural network, the network will have too many degrees of freedom and is trained too long, there is a danger that the network will overfit the data. If there are too few hidden units, the training set cannot be learned. Generally speaking, however, it is better to have too many hidden units than too few. With too few hidden units, the model might not have enough flexibility to capture the nonlinearities in the data; with too many hidden units, the extra weight can be shrunk towards zero if appropriate regularization or pruning, as described below, is used. In typical embodiments, the number of hidden units is somewhere in the

range of 5 to 100, with the number increasing with the number of inputs and number of training cases.

One general approach to determining the number of hidden units to use is to apply a regularization approach. In the regularization approach, a new criterion function is constructed that depends not only on the classical training error, but also on classifier complexity. Specifically, the new criterion function penalizes highly complex models; searching for the minimum in this criterion is to balance error on the training set with error on the training set plus a regularization term, which expresses constraints or desirable properties of solutions:

$$J = J_{\text{pat}} + \lambda J_{\text{reg}}$$

The parameter λ is adjusted to impose the regularization more or less strongly. In other words, larger values for λ will tend to shrink weights towards zero: typically cross-validation with a validation set is used to estimate λ . This validation set can be obtained by setting aside a random subset of the training population. Other forms of penalty can also be used, for example the weight elimination penalty (see, e.g., Hastie *et al.*, 2001, *The Elements of Statistical Learning*, Springer-Verlag, New York).

Another approach to determine the number of hidden units to use is to eliminate - prune - weights that are least needed. In one approach, the weights with the smallest magnitude are eliminated (set to zero). Such magnitude-based pruning can work, but is nonoptimal; sometimes weights with small magnitudes are important for learning and training data. In some embodiments, rather than using a magnitude-based pruning approach, Wald statistics are computed. The fundamental idea in Wald Statistics is that they can be used to estimate the importance of a hidden unit (weight) in a model. Then, hidden units having the least importance are eliminated (by setting their input and output weights to zero). Two algorithms in this regard are the *Optimal Brain Damage* (OBD) and the *Optimal Brain Surgeon* (OBS) algorithms that use second-order approximation to predict how the training error depends upon a weight, and eliminate the weight that leads to the smallest increase in training error.

Optimal Brain Damage and Optimal Brain Surgeon share the same basic approach of training a network to local minimum error at weight w , and then pruning a weight that leads

to the smallest increase in the training error. The predicted functional increase in the error for a change in full weight vector δw is:

$$\Delta J = \left(\frac{\partial J}{\partial w} \right)' \cdot \delta w + \frac{1}{2} \delta w' \cdot \frac{\partial^2 J}{\partial w^2} \cdot \delta w + O(\|\delta w\|^3)$$

where $\frac{\partial^2 J}{\partial w^2}$ is the Hessian matrix. The first term vanishes because we are at a local

- 5 minimum in error; third and higher order terms are ignored. The general solution for minimizing this function given the constraint of deleting one weight is:

$$\delta w = - \frac{w_q}{\left[\mathbf{H}^{-1} \right]_{qq}} \mathbf{H}^{-1} \cdot u_q \text{ and } L_q = \frac{1}{2} - \frac{w_q^2}{\left[\mathbf{H}^{-1} \right]_{qq}}$$

- Here, u_q is the unit vector along the q th direction in weight space and L_q is approximation to the saliency of the weight q - the increase in training error if weight q is pruned and the other weights updated δw . These equations require the inverse of \mathbf{H} . One method to calculate this
- 10 inverse matrix is to start with a small value, $H_0^{-1} = \alpha^{-1} \mathbf{I}$, where α is a small parameter - effectively a weight constant. Next the matrix is updated with each pattern according to

$$\mathbf{H}_{m+1}^{-1} = \mathbf{H}_m^{-1} - \frac{\mathbf{H}_m^{-1} \mathbf{X}_{m+1} \mathbf{X}_{m+1}^T \mathbf{H}_m^{-1}}{\frac{n}{a_m} + \mathbf{X}_{m+1}^T \mathbf{H}_m^{-1} \mathbf{X}_{m+1}}$$

- where the subscripts correspond to the pattern being presented and a_m decreases with m .
- 15 After the full training set has been presented, the inverse Hessian matrix is given by $\mathbf{H}^{-1} = \mathbf{H}_n^{-1}$. In algorithmic form, the Optimal Brain Surgeon method is:

begin initialize n_H, w, θ

train a reasonably large network to minimum error

do compute \mathbf{H}^{-1} by Eqn. 1

- 20 $q^* \leftarrow \arg \min_q w_q^2 / (2 \left[\mathbf{H}^{-1} \right]_{qq})$ (saliency L_q)

$$\mathbf{w} \leftarrow \mathbf{w} - \frac{\mathbf{w}_{q^*}}{[\mathbf{H}^{-1}]_{q^*,q^*}} \mathbf{H}^{-1} \mathbf{e}_{q^*} \quad (\text{saliency } L_q)$$

until $J(\mathbf{w}) > \theta$

return \mathbf{w}

end

5 The Optimal Brain Damage method is computationally simpler because the calculation of the inverse Hessian matrix in line 3 is particularly simple for a diagonal matrix. The above algorithm terminates when the error is greater than a criterion initialized to be θ . Another approach is to change line 6 to terminate when the change in $J(\mathbf{w})$ due to elimination of a weight is greater than some criterion value.

10 In some embodiments, a back-propagation neural network (see, for example Abdi, 1994, "A neural network primer", J. Biol System. 2, 247-283) containing a single hidden layer of ten neurons (ten hidden units) found in EasyNN-Plus version 4.0g software package (Neural Planner Software Inc.) is used. In a specific example, parameter values within the EasyNN-Plus program are set as follows: a learning rate of 0.05, and a momentum of 0.2. In
15 some embodiments in which the EasyNN-Plus version 4.0g software package is used, "outlier" samples are identified by performing twenty independently-seeded trials involving 20,000 learning cycles each.

5.2.4.3. SUPPORT VECTOR MACHINE

20 In some embodiments of the present invention, support vector machines (SVMs) are used to classify subjects using expression profiles of marker genes described in the present invention. General description of SVM can be found in, for example, Cristianini and Shawe-Taylor, 2000, *An Introduction to Support Vector Machines*, Cambridge University Press, Cambridge; Boser *et al.*, 1992, "A training algorithm for optimal margin classifiers, in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, ACM
25 Press, Pittsburgh, PA, pp. 142-152; Vapnik, 1998, *Statistical Learning Theory*, Wiley, New York; Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc.; Hastie, 2001, *The Elements of Statistical Learning*, Springer, New York; and Furey *et al.*, 2000, *Bioinformatics* 16, 906-914. Applications of SVM in biological applications are described in Jaakkola *et al.*, *Proceedings of the 7th International Conference on Intelligent Systems for*

Molecular Biology, AAAI Press, Menlo Park, CA (1999); Brown *et al.*, *Proc. Natl. Acad. Sci.* 97(1):262-67 (2000); Zien *et al.*, *Bioinformatics*, 16(9):799-807 (2000); Furey *et al.*, *Bioinformatics*, 16(10):906-914 (2000)

In one approach, when a SVM is used, the gene expression data is standardized to have mean zero and unit variance and the members of a training population are randomly divided into a training set and a test set. For example, in one embodiment, two thirds of the members of the training population are placed in the training set and one third of the members of the training population are placed in the test set. The expression values for a selected set of genes of the present invention is used to train the SVM. Then the ability for the trained SVM to correctly classify members in the test set is determined. In some embodiments, this computation is performed several times for a given selected set of molecular markers. In each iteration of the computation, the members of the training population are randomly assigned to the training set and the test set. Then, the quality of the combination of molecular markers is taken as the average of each such iteration of the SVM computation.

Support vector machines map a given set of binary labeled training data to a high-dimensional feature space and separate the two classes of data with a maximum margin hyperplane. In general, this hyperplane corresponds to a nonlinear decision boundary in the input space. Let $X \in R_0 \subseteq \mathcal{R}^n$ be the input vectors, $y \in \{-1, +1\}$ be the labels, and $\phi: R_0 \rightarrow F$ be the mapping from input space to feature space. Then the SVM learning algorithm finds a hyperplane (w, b) such that the quantity

$$\gamma = \min_i y_i \{ \langle w, \phi(X_i) \rangle - b \}$$

is maximized, where the vector w has the same dimensionality as F , b is a real number, and γ is called the *margin*. The corresponding decision function is then

$$f(X) = \text{sign}(\langle w, \phi(X) \rangle - b)$$

This minimum occurs when

$$w = \sum_i \alpha_i y_i \phi(X_i)$$

where $\{\alpha_i\}$ are positive real numbers that maximize

$$\sum_i \alpha_i - \sum_{ij} \alpha_i \alpha_j y_i y_j \langle \phi(X_i), \phi(X_j) \rangle$$

subject to

$$\sum_i \alpha_i y_i = 0, \alpha_i > 0$$

5 The decision function can equivalently be expressed as

$$f(\mathbf{X}) = \text{sign}(\sum_i \alpha_i y_i \langle \phi(X_i), \phi(\mathbf{X}) \rangle - b)$$

From this equation it can be seen that the α_i associated with the training point \mathbf{X}_i expresses the strength with which that point is embedded in the final decision function. A remarkable property of this alternative representation is that only a subset of the points will be associated with a non-zero α_i . These points are called *support vectors* and are the points that lie closest to the separating hyperplane. The sparseness of the α vector has several computational and learning theoretic consequences. It is important to note that neither the learning algorithm nor the decision function needs to represent explicitly the image of points in the feature space, $\phi(\mathbf{X}_i)$, since both use only the dot products between such images, $\langle \phi(X_i), \phi(X_j) \rangle$. Hence, if one were given a function $K(\mathbf{X}, \mathbf{Y}) = \langle \phi(\mathbf{X}), \phi(\mathbf{Y}) \rangle$, one could learn and use the maximum margin hyperplane in the feature space without ever explicitly performing the mapping. For each continuous positive definite function $K(\mathbf{X}, \mathbf{Y})$ there exists a mapping ϕ such that $K(\mathbf{X}, \mathbf{Y}) = \langle \phi(\mathbf{X}), \phi(\mathbf{Y}) \rangle$ for all $\mathbf{X}, \mathbf{Y} \in R_0$ (Mercer's Theorem). The function $K(\mathbf{X}, \mathbf{Y})$ is called the *kernel function*. The use of a kernel function allows the support vector machine to operate efficiently in a nonlinear high-dimensional feature spaces without being adversely affected by the dimensionality of that space. Indeed, it is possible to work with feature spaces of infinite dimension. Moreover, Mercer's theorem makes it possible to learn in the feature space without even knowing ϕ and F . The matrix $K_{ij} = \langle \phi(\mathbf{X}_i), \phi(\mathbf{X}_j) \rangle$ is called the *kernel matrix*. Finally, note that the learning algorithm is a quadratic optimization problem that has only a global optimum. The absence of local minima is a significant difference from standard pattern recognition techniques such as neural

networks. For moderate sample sizes, the optimization problem can be solved with simple gradient descent techniques. In the presence of noise, the standard maximum margin algorithm described above can be subject to overfitting, and more sophisticated techniques should be used. This problem arises because the maximum margin algorithm always finds a perfectly consistent hypothesis and does not tolerate training error. Sometimes, however, it is necessary to trade some training accuracy for better predictive power. The need for tolerating training error has led to the development the soft-margin and the margin-distribution classifiers. One of these techniques replaces the kernel matrix in the training phase as follows:

$$K \leftarrow K + \lambda I$$

while still using the standard kernel function in the decision phase. By tuning λ , one can control the training error, and it is possible to prove that the risk of misclassifying unseen points can be decreased with a suitable choice of λ .

If instead of controlling the overall training error one wants to control the trade-off between false positives and false negatives, it is possible to modify K as follows:

$$K \leftarrow K + \lambda D$$

where D is a diagonal matrix whose entries are either d^+ or d^- , in locations corresponding to positive and negative examples. It is possible to prove that this technique is equivalent to controlling the size of the α_i in a way that depends on the size of the class, introducing a bias for larger α_i in the class with smaller d . This in turn corresponds to an asymmetric margin; i.e., the class with smaller d will be kept further away from the decision boundary. In some cases, the extreme imbalance of the two classes, along with the presence of noise, creates a situation in which points from the minority class can be easily mistaken for mislabelled points. Enforcing a strong bias against training errors in the minority class provides protection against such errors and forces the SVM to make the positive examples support vectors. Thus, choosing $d^+ = \frac{1}{n^+}$ and $d^- = \frac{1}{n^-}$ provides a heuristic way to automatically adjust the relative importance of the two classes, based on their respective cardinalities. This technique effectively controls the trade-off between sensitivity and specificity.

In the present invention, a linear kernel can be used. The similarity between two marker profiles X and Y can be the dot product $X \cdot Y$. In one embodiment, the kernel is

$$K(X, Y) = X \cdot Y + 1$$

In another embodiment, a kernel of degree d is used

$$K(X, Y) = (X \cdot Y + 1)^d, \text{ where } d \text{ can be either } 2, 3, \dots$$

In still another embodiment, a Gaussian kernel is used

$$K(X, Y) = \exp\left(-\frac{\|X - Y\|^2}{2\sigma^2}\right)$$

where σ is the width of the Gaussian.

5.2.4.4. LOGISTIC REGRESSION

In some embodiments, the classifier is based on a regression model, preferably a logistic regression model. Such a regression model includes a coefficient for each of the molecular markers in a selected set of molecular markers of the invention. In such embodiments, the coefficients for the regression model are computed using, for example, a maximum likelihood approach. In particular embodiments, molecular marker data from two different clinical groups, e.g., good or poor prognosis, is used and the dependent variable is the clinical status of the patient for which molecular marker characteristic data are from.

Some embodiments of the present invention provide generalizations of the logistic regression model that handle multicategory (polychotomous) responses. Such embodiments can be used to discriminate an organism into one or three or more clinical groups, e.g., good, intermediate, and poor prognosis. Such regression models use multicategory logit models that simultaneously refer to all pairs of categories, and describe the odds of response in one category instead of another. Once the model specifies logits for a certain (J-1) pairs of categories, the rest are redundant. See, for example, Agresti, *An Introduction to Categorical Data Analysis*, John Wiley & Sons, Inc., 1996, New York, Chapter 8, which is hereby incorporated by reference.

5.2.4.5. DISCRIMINANT ANALYSIS

Linear discriminant analysis (LDA) attempts to classify a subject into one of two categories based on certain object properties. In other words, LDA tests whether object attributes measured in an experiment predict categorization of the objects. LDA typically requires continuous independent variables and a dichotomous categorical dependent variable. In the present invention, the expression values for the selected set of molecular markers of the

invention across a subset of the training population serve as the requisite continuous independent variables. The clinical group classification of each of the members of the training population serves as the dichotomous categorical dependent variable.

LDA seeks the linear combination of variables that maximizes the ratio of between-group variance and within-group variance by using the grouping information. Implicitly, the linear weights used by LDA depend on how the expression of a molecular marker across the training set separates in the two groups (e.g., a group that has good prognosis and a group that have poor prognosis) and how this gene expression correlates with the expression of other genes. In some embodiments, LDA is applied to the data matrix of the N members in a training sample by K genes in a combination of genes described in the present invention. Then, the linear discriminant of each member of the training population is plotted. Ideally, those members of the training population representing a first subgroup (e.g. those subjects that good prognosis) will cluster into one range of linear discriminant values (e.g., negative) and those member of the training population representing a second subgroup (e.g. those subjects that have poor prognosis) will cluster into a second range of linear discriminant values (e.g., positive). The LDA is considered more successful when the separation between the clusters of discriminant values is larger. For more information on linear discriminant analysis, see Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc; and Hastie, 2001, *The Elements of Statistical Learning*, Springer, New York; Venables & Ripley, 1997, *Modern Applied Statistics with s-plus*, Springer, New York.

Quadratic discriminant analysis (QDA) takes the same input parameters and returns the same results as LDA. QDA uses quadratic equations, rather than linear equations, to produce results. LDA and QDA are interchangeable, and which to use is a matter of preference and/or availability of software to support the analysis. Logistic regression takes the same input parameters and returns the same results as LDA and QDA.

5.2.4.6. DECISION TREES

In some embodiments of the present invention, decision trees are used to classify patients using expression data for a selected set of molecular markers of the invention. Decision tree algorithms belong to the class of supervised learning algorithms. The aim of a decision tree is to induce a classifier (a tree) from real-world example data. This tree can be used to classify unseen examples which have not been used to derive the decision tree.

A decision tree is derived from training data. An example contains values for the different attributes and what class the example belongs. In one embodiment, the training data is expression data for a combination of genes described in the present invention across the training population.

- 5 The following algorithm describes a decision tree derivation:

Tree(Examples,Class,Attributes)

 Create a root node

 If all Examples have the same Class value, give the root this label

 Else if Attributes is empty label the root according to the most common value

- 10 Else begin

 Calculate the information gain for each attribute

 Select the attribute A with highest information gain and make this the root
attribute

 For each possible value, v, of this attribute

- 15 Add a new branch below the root, corresponding to $A = v$

 Let Examples(v) be those examples with $A = v$

 If Examples(v) is empty, make the new branch a leaf node labeled with
the most common value among Examples

 Else let the new branch be the tree created by

- 20 Tree(Examples(v),Class,Attributes - {A})

 end

A more detailed description of the calculation of information gain is shown in the following. If the possible classes v_i of the examples have probabilities $P(v_i)$ then the information content I of the actual answer is given by:

$$I(P(v_1), \dots, P(v_n)) = - \sum_{i=1}^n P(v_i) \log_2 P(v_i)$$

- 5 The I -value shows how much information we need in order to be able to describe the outcome of a classification for the specific dataset used. Supposing that the dataset contains p positive (e.g. has poor prognosis) and n negative (e.g. has good prognosis) examples (e.g. individuals), the information contained in a correct answer is:

$$I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n}$$

- 10 where \log_2 is the logarithm using base two. By testing single attributes the amount of information needed to make a correct classification can be reduced. The remainder for a specific attribute A (e.g. a gene) shows how much the information that is needed can be reduced.

$$\text{Remainder}(A) = \sum_{i=1}^v \frac{p_i + n_i}{p+n} I\left(\frac{p_i}{p_i + n_i}, \frac{n_i}{p_i + n_i}\right)$$

- 15 " v " is the number of unique attribute values for attribute A in a certain dataset, " i " is a certain attribute value, " p_i " is the number of examples for attribute A where the classification is positive (e.g. cancer), " n_i " is the number of examples for attribute A where the classification is negative (e.g. healthy).

- 20 The information gain of a specific attribute A is calculated as the difference between the information content for the classes and the remainder of attribute A :

$$\text{Gain}(A) = I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) - \text{Remainder}(A)$$

The information gain is used to evaluate how important the different attributes are for the classification (how well they split up the examples), and the attribute with the highest information.

- 5 In general there are a number of different decision tree algorithms, many of which are described in Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc. Decision tree algorithms often require consideration of feature processing, impurity measure, stopping criterion, and pruning. Specific decision tree algorithms include, but are not limited to classification and regression trees (CART), multivariate decision trees, ID3, and C4.5.

- 10 In one approach, when an exemplary embodiment of a decision tree is used, the gene expression data for a selected set of molecular markers of the invention across a training population is standardized to have mean zero and unit variance. The members of the training population are randomly divided into a training set and a test set. For example, in one embodiment, two thirds of the members of the training population are placed in the training set and one third of the members of the training population are placed in the test set. The
15 expression values for a select combination of genes described in the present invention is used to construct the decision tree. Then, the ability for the decision tree to correctly classify members in the test set is determined. In some embodiments, this computation is performed several times for a given combination of molecular markers. In each iteration of the computation, the members of the training population are randomly assigned to the training set
20 and the test set. Then, the quality of the combination of molecular markers is taken as the average of each such iteration of the decision tree computation.

5.2.4.7. CLUSTERING

- In some embodiments, the expression values for a selected set of molecular markers of the invention are used to cluster a training set. For example, consider the case in which ten
25 genes described in the present invention are used. Each member *m* of the training population will have expression values for each of the ten genes. Such values from a member *m* in the training population define the vector:

X_{1m}	X_{2m}	X_{3m}	X_{4m}	X_{5m}	X_{6m}	X_{7m}	X_{8m}	X_{9m}	X_{10m}
----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------

where X_{im} is the expression level of the i^{th} gene in organism m . If there are m organisms in the training set, selection of i genes will define m vectors. Note that the methods of the present invention do not require that each the expression value of every single gene used in the vectors be represented in every single vector m . In other words, data from a subject in which one of the i^{th} genes is not found can still be used for clustering. In such instances, the missing expression value is assigned either a "zero" or some other normalized value. In some embodiments, prior to clustering, the gene expression values are normalized to have a mean value of zero and unit variance.

Those members of the training population that exhibit similar expression patterns across the training group will tend to cluster together. A particular combination of genes of the present invention is considered to be a good classifier in this aspect of the invention when the vectors cluster into the trait groups found in the training population. For instance, if the training population includes patients with good or poor prognosis, a clustering classifier will cluster the population into two groups, with each group uniquely representing either a good prognosis or a poor prognosis.

Clustering is described on pages 211-256 of Duda and Hart, *Pattern Classification and Scene Analysis*, 1973, John Wiley & Sons, Inc., New York. As described in Section 6.7 of Duda, the clustering problem is described as one of finding natural groupings in a dataset.

To identify natural groupings, two issues are addressed. First, a way to measure similarity (or dissimilarity) between two samples is determined. This metric (similarity measure) is used to ensure that the samples in one cluster are more like one another than they are to samples in other clusters. Second, a mechanism for partitioning the data into clusters using the similarity measure is determined.

Similarity measures are discussed in Section 6.7 of Duda, where it is stated that one way to begin a clustering investigation is to define a distance function and to compute the matrix of distances between all pairs of samples in a dataset. If distance is a good measure of similarity, then the distance between samples in the same cluster will be significantly less than the distance between samples in different clusters. However, as stated on page 215 of Duda, clustering does not require the use of a distance metric. For example, a nonmetric similarity function $s(x, x')$ can be used to compare two vectors x and x' . Conventionally, $s(x,$

x') is a symmetric function whose value is large when x and x' are somehow "similar". An example of a nonmetric similarity function $s(x, x')$ is provided on page 216 of Duda.

Once a method for measuring "similarity" or "dissimilarity" between points in a dataset has been selected, clustering requires a criterion function that measures the clustering quality of any partition of the data. Partitions of the data set that extremize the criterion function are used to cluster the data. See page 217 of Duda. Criterion functions are discussed in Section 6.8 of Duda.

More recently, Duda *et al.*, Pattern Classification, 2nd edition, John Wiley & Sons, Inc. New York, has been published. Pages 537-563 describe clustering in detail. More information on clustering techniques can be found in Kaufman and Rousseeuw, 1990, *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York, NY; Everitt, 1993, *Cluster analysis* (3d ed.), Wiley, New York, NY; and Backer, 1995, *Computer-Assisted Reasoning in Cluster Analysis*, Prentice Hall, Upper Saddle River, New Jersey. Particular exemplary clustering techniques that can be used in the present invention include, but are not limited to, hierarchical clustering (agglomerative clustering using nearest-neighbor algorithm, farthest-neighbor algorithm, the average linkage algorithm, the centroid algorithm, or the sum-of-squares algorithm), k-means clustering, fuzzy k-means clustering algorithm, and Jarvis-Patrick clustering.

5.2.4.8. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) has been proposed to analyze gene expression data. Principal component analysis is a classical technique to reduce the dimensionality of a data set by transforming the data to a new set of variable (principal components) that summarize the features of the data. See, for example, Jolliffe, 1986, *Principal Component Analysis*, Springer, New York. Principal components (PCs) are uncorrelated and are ordered such that the k^{th} PC has the k^{th} largest variance among PCs. The k^{th} PC can be interpreted as the direction that maximizes the variation of the projections of the data points such that it is orthogonal to the first $k - 1$ PCs. The first few PCs capture most of the variation in the data set. In contrast, the last few PCs are often assumed to capture only the residual 'noise' in the data.

PCA can also be used to create a classifier in accordance with the present invention. In such an approach, vectors for a selected set of molecular markers of the invention can be

constructed in the same manner described for clustering above. In fact, the set of vectors, where each vector represents the expression values for the select genes from a particular member of the training population, can be considered a matrix. In some embodiments, this matrix is represented in a Free-Wilson method of qualitative binary description of monomers (Kubinyi, 1990, *3D QSAR in drug design theory methods and applications*, Pergamon Press, Oxford, pp 589-638), and distributed in a maximally compressed space using PCA so that the first principal component (PC) captures the largest amount of variance information possible, the second principal component (PC) captures the second largest amount of all variance information, and so forth until all variance information in the matrix has been accounted for.

- 10 Then, each of the vectors (where each vector represents a member of the training population) is plotted. Many different types of plots are possible. In some embodiments, a one-dimensional plot is made. In this one-dimensional plot, the value for the first principal component from each of the members of the training population is plotted. In this form of plot, the expectation is that members of a first group (e.g. good prognosis patients) will
- 15 cluster in one range of first principal component values and members of a second group (e.g., poor prognosis patients) will cluster in a second range of first principal component values.

- In one example, the training population comprises two groups: a good prognosis group and a poor prognosis group. The first principal component is computed using the molecular marker expression values for the select genes of the present invention across the
- 20 entire training population data set. Then, each member of the training set is plotted as a function of the value for the first principal component. In this example, those members of the training population in which the first principal component is positive are the good prognosis patients and those members of the training population in which the first principal component is negative are poor prognosis patients.

- 25 In some embodiments, the members of the training population are plotted against more than one principal component. For example, in some embodiments, the members of the training population are plotted on a two-dimensional plot in which the first dimension is the first principal component and the second dimension is the second principal component. In such a two-dimensional plot, the expectation is that members of each subgroup represented in
- 30 the training population will cluster into discrete groups. For example, a first cluster of members in the two-dimensional plot will represent subjects with good prognosis, a second

cluster of members in the two-dimensional plot will represent subjects with poor prognosis, and so forth.

- In some embodiments, the members of the training population are plotted against more than two principal components and a determination is made as to whether the members of the training population are clustering into groups that each uniquely represents a subgroup found in the training population. In some embodiments, principal component analysis is performed by using the *R mva* package (Anderson, 1973, *Cluster Analysis for applications*, Academic Press, New York 1973; Gordon, *Classification*, Second Edition, Chapman and Hall, CRC, 1999.). Principal component analysis is further described in Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc.

5.2.4.9. NEAREST NEIGHBOR CLASSIFIER ANALYSIS

- Nearest neighbor classifiers are memory-based and require no model to be fit. Given a query point x_0 , the k training points $x_{(1)}, \dots, x_{(k)}$ closest in distance to x_0 are identified and then the point x_0 is classified using the k nearest neighbors. Ties can be broken at random. In some embodiments, Euclidean distance in feature space is used to determine distance as:

$$d_{(i)} = \|x_{(i)} - x_0\|.$$

- Typically, when the nearest neighbor algorithm is used, the expression data used to compute the linear discriminant is standardized to have mean zero and variance 1. In the present invention, the members of the training population are randomly divided into a training set and a test set. For example, in one embodiment, two thirds of the members of the training population are placed in the training set and one third of the members of the training population are placed in the test set. Profiles of a selected set of molecular markers of the invention represents the feature space into which members of the test set are plotted. Next, the ability of the training set to correctly characterize the members of the test set is computed. In some embodiments, nearest neighbor computation is performed several times for a given combination of genes of the present invention. In each iteration of the computation, the members of the training population are randomly assigned to the training set and the test set. Then, the quality of the combination of genes is taken as the average of each such iteration of the nearest neighbor computation.

The nearest neighbor rule can be refined to deal with issues of unequal class priors, differential misclassification costs, and feature selection. Many of these refinements involve some form of weighted voting for the neighbors. For more information on nearest neighbor analysis, see Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc; and Hastie, 2001, *The Elements of Statistical Learning*, Springer, New York.

5.2.4.10. EVOLUTIONARY METHODS

Inspired by the process of biological evolution, evolutionary methods of classifier design employ a stochastic search for an optimal classifier. In broad overview, such methods create several classifiers - a population - from measurements of gene products of the present invention. Each classifier varies somewhat from the other. Next, the classifiers are scored on expression data across the training population. In keeping with the analogy with biological evolution, the resulting (scalar) score is sometimes called the fitness. The classifiers are ranked according to their score and the best classifiers are retained (some portion of the total population of classifiers). Again, in keeping with biological terminology, this is called survival of the fittest. The classifiers are stochastically altered in the next generation - the children or offspring. Some offspring classifiers will have higher scores than their parent in the previous generation, some will have lower scores. The overall process is then repeated for the subsequent generation: The classifiers are scored and the best ones are retained, randomly altered to give yet another generation, and so on. In part, because of the ranking, each generation has, on average, a slightly higher score than the previous one. The process is halted when the single best classifier in a generation has a score that exceeds a desired criterion value. More information on evolutionary methods is found in, for example, Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc.

5.2.4.11. BAGGING, BOOSTING AND THE RANDOM SUBSPACE METHOD

Bagging, boosting and the random subspace method are combining techniques that can be used to improve weak classifiers. These techniques are designed for, and usually applied to, decision trees. In addition, Skurichina and Duin provide evidence to suggest that such techniques can also be useful in linear discriminant analysis.

In bagging, one samples the training set, generating random independent bootstrap replicates, constructs the classifier on each of these, and aggregates them by a simple majority vote in the final decision rule. See, for example, Breiman, 1996, *Machine Learning*

24, 123-140; and Efron & Tibshirani, *An Introduction to Bootstrap*, Chapman & Hall, New York, 1993.

In boosting, classifiers are constructed on weighted versions of the training set, which are dependent on previous classification results. Initially, all objects have equal weights, and the first classifier is constructed on this data set. Then, weights are changed according to the performance of the classifier. Erroneously classified objects (molecular markers in the data set) get larger weights, and the next classifier is boosted on the reweighted training set. In this way, a sequence of training sets and classifiers is obtained, which is then combined by simple majority voting or by weighted majority voting in the final decision. See, for example, Freund & Schapire, "Experiments with a new boosting algorithm," Proceedings 13th International Conference on Machine Learning, 1996, 148-156.

To illustrate boosting, consider the case where there are two phenotypic groups exhibited by the population under study, phenotype 1 (e.g., good prognosis patients), and phenotype 2 (e.g., poor prognosis patients). Given a vector of molecular markers X , a classifier $G(X)$ produces a prediction taking one of the type values in the two value set: {phenotype 1, phenotype 2}. The error rate on the training sample is

$$\overline{\text{err}} = \frac{1}{N} \sum_{i=1}^N I(y_i \neq G(x_i))$$

where N is the number of subjects in the training set (the sum total of the subjects that have either phenotype 1 or phenotype 2).

A weak classifier is one whose error rate is only slightly better than random guessing. In the boosting algorithm, the weak classification algorithm is repeatedly applied to modified versions of the data, thereby producing a sequence of weak classifiers $G_m(x)$, $m = 1, 2, \dots, M$. The predictions from all of the classifiers in this sequence are then combined through a weighted majority vote to produce the final prediction:

$$G(x) = \text{sign} \left(\sum_{m=1}^M \alpha_m G_m(x) \right)$$

Here $\alpha_1, \alpha_2, \dots, \alpha_M$ are computed by the boosting algorithm and their purpose is to weigh the contribution of each respective $G_m(x)$. Their effect is to give higher influence to the more accurate classifiers in the sequence.

5 The data modifications at each boosting step consist of applying weights w_1, w_2, \dots, w_n to each of the training observations (x_i, y_i) , $i = 1, 2, \dots, N$. Initially all the weights are set to $w_i = 1/N$, so that the first step simply trains the classifier on the data in the usual manner. For each successive iteration $m = 2, 3, \dots, M$ the observation weights are individually modified and the classification algorithm is reapplied to the weighted observations. At stem
10 m , those observations that were misclassified by the classifier $G_{m-1}(x)$ induced at the previous step have their weights increased, whereas the weights are decreased for those that were classified correctly. Thus as iterations proceed, observations that are difficult to correctly classify receive ever-increasing influence. Each successive classifier is thereby forced to concentrate on those training observations that are missed by previous ones in the sequence.

15 The exemplary boosting algorithm is summarized as follows:

1. Initialize the observation weights $w_i = 1/N$, $i = 1, 2, \dots, N$.

2. For $m = 1$ to M :

20 (a) Fit a classifier $G_m(x)$ to the training set using weights w_i .

(b) Compute

$$\text{err}_m = \frac{\sum_{i=1}^N w_i I(y_i \neq G_m(x_i))}{\sum_{i=1}^N w_i}$$

(c) Compute $\alpha_m = \log((1 - \text{err}_m) / \text{err}_m)$.

(d) Set $w_i \leftarrow w_i \cdot \exp[\alpha_m \cdot I(y_i \neq G_m(x_i))]$, $i = 1, 2, \dots, N$.

3. Output $G(x) = \text{sign} \left[\sum_{m=1}^M \alpha_m G_m(x) \right]$

In the algorithm, the current classifier $G_m(x)$ is induced on the weighted observations at line

- 5 2a. The resulting weighted error rate is computed at line 2b. Line 2c calculates the weight α_m given to $G_m(x)$ in producing the final classifier $G(x)$ (line 3). The individual weights of each of the observations are updated for the next iteration at line 2d. Observations misclassified by $G_m(x)$ have their weights scaled by a factor $\exp(\alpha_m)$, increasing their relative influence for inducing the next classifier $G_{m+1}(x)$ in the sequence. In some embodiments, modifications of
- 10 the Freund and Schapire, 1997, Journal of Computer and System Sciences 55, pp. 119-139, boosting method are used. See, for example, Hasti *et al.*, *The Elements of Statistical Learning*, 2001, Springer, New York, Chapter 10. In some embodiments, boosting or adaptive boosting methods are used.

- In some embodiments, modifications of Freund and Schapire, 1997, Journal of
- 15 Computer and System Sciences 55, pp. 119-139, are used. For example, in some embodiments, feature preselection is performed using a technique such as the nonparametric scoring methods of Park *et al.*, 2002, Pac. Symp. Biocomput. 6, 52-63. Feature preselection is a form of dimensionality reduction in which the genes that discriminate between classifications the best are selected for use in the classifier. Then, the LogitBoost procedure
- 20 introduced by Friedman *et al.*, 2000, Ann Stat 28, 337-407 is used rather than the boosting procedure of Freund and Schapire. In some embodiments, the boosting and other classification methods of Ben-Dor *et al.*, 2000, Journal of Computational Biology 7, 559-583 are used in the present invention. In some embodiments, the boosting and other classification methods of Freund and Schapire, 1997, Journal of Computer and System Sciences 55, 119-
- 25 139, are used.

- In the random subspace method, classifiers are constructed in random subspaces of the data feature space. These classifiers are usually combined by simple majority voting in the final decision rule. See, for example, Ho, "The Random subspace method for constructing decision forests," IEEE Trans Pattern Analysis and Machine Intelligence, 1998;
- 30 20(8): 832-844.

5.2.4.12. OTHER ALGORITHMS

The pattern classification and statistical techniques described above are merely examples of the types of models that can be used to construct a model for classification. Moreover, combinations of the techniques described above can be used. Some combinations, such as the use of the combination of decision trees and boosting, have been described. However, many other combinations are possible. In addition, in other techniques in the art such as Projection Pursuit and Weighted Voting can be used to construct a classifier.

5.3. DETERMINATION OF MARKER GENE EXPRESSION LEVELS

5.3.1. METHODS

The expression levels of the marker genes in a sample may be determined by any means known in the art. The expression level may be determined by isolating and determining the level (*i.e.*, amount) of nucleic acid transcribed from each marker gene. Alternatively, or additionally, the level of specific proteins encoded by a marker gene may be determined.

The level of expression of specific marker genes can be accomplished by determining the amount of mRNA, or polynucleotides derived therefrom, present in a sample. Any method for determining RNA levels can be used. For example, RNA is isolated from a sample and separated on an agarose gel. The separated RNA is then transferred to a solid support, such as a filter. Nucleic acid probes representing one or more markers are then hybridized to the filter by northern hybridization, and the amount of marker-derived RNA is determined. Such determination can be visual, or machine-aided, for example, by use of a densitometer. Another method of determining RNA levels is by use of a dot-blot or a slot-blot. In this method, RNA, or nucleic acid derived therefrom, from a sample is labeled. The RNA or nucleic acid derived therefrom is then hybridized to a filter containing oligonucleotides derived from one or more marker genes, wherein the oligonucleotides are placed upon the filter at discrete, easily-identifiable locations. Hybridization, or lack thereof, of the labeled RNA to the filter-bound oligonucleotides is determined visually or by densitometer. Polynucleotides can be labeled using a radiolabel or a fluorescent (*i.e.*, visible) label.

These examples are not intended to be limiting; other methods of determining RNA abundance are known in the art.

The level of expression of particular marker genes may also be assessed by determining the level of the specific protein expressed from the marker genes. This can be accomplished, for example, by separation of proteins from a sample on a polyacrylamide gel, followed by identification of specific marker-derived proteins using antibodies in a western blot. Alternatively, proteins can be separated by two-dimensional gel electrophoresis systems. Two-dimensional gel electrophoresis is well-known in the art and typically involves isoelectric focusing along a first dimension followed by SDS-PAGE electrophoresis along a second dimension. See, e.g., Hames *et al.*, 1990, GEL ELECTROPHORESIS OF PROTEINS: A PRACTICAL APPROACH, IRL Press, New York; Shevchenko *et al.*, *Proc. Nat'l Acad. Sci. USA* 93:1440-1445 (1996); Sagliocco *et al.*, *Yeast* 12:1519-1533 (1996); Lander, *Science* 274:536-539 (1996). The resulting electropherograms can be analyzed by numerous techniques, including mass spectrometric techniques, western blotting and immunoblot analysis using polyclonal and monoclonal antibodies.

Alternatively, marker-derived protein levels can be determined by constructing an antibody microarray in which binding sites comprise immobilized, preferably monoclonal, antibodies specific to a plurality of protein species encoded by the cell genome. Preferably, antibodies are present for a substantial fraction of the marker-derived proteins of interest. Methods for making monoclonal antibodies are well known (see, e.g., Harlow and Lane, 1988, ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor, New York, which is incorporated in its entirety for all purposes). In one embodiment, monoclonal antibodies are raised against synthetic peptide fragments designed based on genomic sequence of the cell. With such an antibody array, proteins from the cell are contacted to the array, and their binding is assayed with assays known in the art. Generally, the expression, and the level of expression, of proteins of diagnostic or prognostic interest can be detected through immunohistochemical staining of tissue slices or sections.

Finally, expression of marker genes in a number of tissue specimens may be characterized using a "tissue array" (Kononen *et al.*, *Nat. Med* 4(7):844-7 (1998)). In a tissue array, multiple tissue samples are assessed on the same microarray. The arrays allow *in situ* detection of RNA and protein levels; consecutive sections allow the analysis of multiple samples simultaneously.

5.3.2. MICROARRAYS

In preferred embodiments, polynucleotide microarrays are used to measure expression so that the expression status of each of the markers above is assessed simultaneously.

Generally, microarrays according to the invention comprise a plurality of markers

informative for prognosis, or outcome determination, for a particular disease or condition,

- 5 and, in particular, for individuals having specific combinations of genotypic or phenotypic characteristics of the disease or condition (*i.e.*, that are prognosis-informative for a particular patient subset).

The microarrays of the invention preferably comprise at least 2, 3, 4, 5, 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200 or more of markers, or all of the markers, or any

- 10 combination of markers, identified as prognosis-informative within a patient subset. The actual number of informative markers the microarray comprises will vary depending upon the particular condition of interest, the number of markers identified, and, optionally, the number of informative markers found to result in the least Type I error, Type II error, or Type I and Type II error in determination of prognosis. As used herein, "Type I error" means a false
15 positive and "Type II error" means a false negative; in the example of prognosis of breast cancer, Type I error is the mis-characterization of an individual with a good prognosis as having a poor prognosis, and Type II error is the mis-characterization of an individual with a poor prognosis as having a good prognosis.

In specific embodiments, the invention provides polynucleotide arrays in which the
20 prognosis markers identified for a particular patient subset comprise at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the probes on said array. In another specific embodiment, the microarray comprises a plurality of probes, wherein said plurality of probes comprise probes complementary and hybridizable to at least 75% of the prognosis-informative markers identified for a particular patient subset. Microarrays of the invention, of course, may

- 25 comprise probes complementary and hybridizable to prognosis-informative markers for a plurality of the patient subsets, or for each patient subset, identified for a particular condition. In another embodiment, therefore, the microarray of the invention comprises a plurality of probes complementary and hybridizable to at least 75% of the prognosis-informative markers identified for each patient subset identified for the condition of interest, and wherein said
30 probes, in total, are at least 50% of the probes on said microarray.

In yet another specific embodiment, microarrays that are used in the methods disclosed herein optionally comprise markers additional to at least some of the markers

identified by the methods disclosed elsewhere herein. For example, in a specific embodiment, the microarray is a screening or scanning array as described in Altschuler *et al.*, International Publication WO 02/18646, published March 7, 2002 and Scherer *et al.*, International Publication WO 02/16650, published February 28, 2002. The scanning and screening arrays comprise regularly-spaced, positionally-addressable probes derived from genomic nucleic acid sequence, both expressed and unexpressed. Such arrays may comprise probes corresponding to a subset of, or all of, the markers identified for the patient subset(s) for the condition of interest, and can be used to monitor marker expression in the same way as a microarray containing only prognosis-informative markers otherwise identified.

In yet another specific embodiment, the microarray is a commercially-available cDNA microarray that comprises at least five markers identified by the methods described herein. Preferably, a commercially-available cDNA microarray comprises all of the markers identified by the methods described herein as being informative for a patient subset for a particular condition. However, such a microarray may comprise at least 5, 10, 15 or 25 of such markers, up to the maximum number of markers identified.

In an embodiment specific to breast cancer, the invention provides for oligonucleotide or cDNA arrays comprising probes hybridizable to the genes corresponding to each of the marker sets described above (*i.e.*, markers informative for ER⁻, sporadic individuals, markers informative for ER⁻, *BRCA1* individuals, markers informative for ER⁺, ER/AGE high individuals, markers informative for ER⁺, ER/AGE low, LN⁺ individuals, and markers informative for ER⁺, ER/AGE low, LN⁻ individuals, as shown in Tables 1-5). Any of the microarrays described herein may be provided in a sealed container in a kit.

The invention provides microarrays containing probes useful for the prognosis of any breast cancer patient, or for breast cancer patients classified into one of a plurality of patient subsets. In particular, the invention provides polynucleotide arrays comprising probes to a subset or subsets of at least 5, 10, 15, 20, 25 or more of the genetic markers, or up to the full set of markers, in any of Tables 1-5, which distinguish between patients with good and poor prognosis. In certain embodiments, therefore, the invention provides microarrays comprising probes for a plurality of the genes for which markers are listed in Tables 1, 2, 3, 4 or 5. In a specific embodiment, the microarray of the invention comprises 1, 2, 3, 4, 5 or 10 of the markers in Table 1, at least five of the markers in Table 2; 1, 2, 3, 4, 5 or 10 of the markers in Table 3; 1, 2, 3, 4, 5 or 10 of the markers in Table 4; or 1, 2, 3, 4, 5 or 10 of the markers in

Table 1. In other embodiments, the microarray comprises probes for 1, 2, 3, 4, 5, or 10 of the markers shown in any two, three or four of Tables 1-5, or all of Tables 1-5. In other embodiments, the microarray of the invention contains each of the markers in Table 1, Table 2, Table 3, Table 4, or Table 5. In another embodiment, the microarray contains all of the markers shown in Tables 1-5. In specific embodiments, the array comprises probes derived only from the markers listed in Table 1, Table 2, Table 3, Table 4, or Table 5; probes derived from any two of Tables 1-5; any three of Tables 1-5; any four of Tables 1-5; or all of Tables 1-5.

In other embodiments, the array comprises a plurality of probes derived from markers listed in any of Tables 1-5 in combination with a plurality of other probes, derived from markers not listed in any of Tables 1-5, that are identified as informative for the prognosis of breast cancer.

In specific embodiments, the invention provides polynucleotide arrays in which the breast cancer prognosis markers described herein in Tables 1, 2, 3, 4 and/or 5 comprise at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the probes on said array. In another specific embodiment, the microarray comprises a plurality of probes, wherein said plurality of probes comprise probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 1; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 2; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 3; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 4; and probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 5, wherein said probes, in total, are at least 50% of the probes on said microarray.

In yet another specific embodiment, microarrays that are used in the methods disclosed herein optionally comprise markers additional to at least some of the markers listed in Tables 1-5. For example, in a specific embodiment, the microarray is a screening or scanning array as described in Altschuler *et al.*, International Publication WO 02/18646, published March 7, 2002 and Scherer *et al.*, International Publication WO 02/16650, published February 28, 2002. The scanning and screening arrays comprise regularly-spaced, positionally-addressable probes derived from genomic nucleic acid sequence, both expressed and unexpressed. Such arrays may comprise probes corresponding to a subset of, or all of,

the markers listed in Tables 1-5, or a subset thereof as described above, and can be used to monitor marker expression in the same way as a microarray containing only markers listed in Tables 1-5.

In yet another specific embodiment, the microarray is a commercially-available cDNA microarray that comprises at least five of the markers listed in Tables 1-5. Preferably, a commercially-available cDNA microarray comprises all of the markers listed in Tables 1-5. However, such a microarray may comprise at least 5, 10, 15 or 25 of the markers in any of Tables 1-5, up to the maximum number of markers in a Table, and may comprise all of the markers in any one of Tables 1-5, and a subset of another of Tables 1-5, or subsets of each as described above. In a specific embodiment of the microarrays used in the methods disclosed herein, the markers that are all or a portion of Tables 1-5 make up at least 50%, 60%, 70%, 80%, 90%, 95% or 98% of the probes on the microarray.

General methods pertaining to the construction of microarrays comprising the marker sets and/or subsets above are described in the following sections.

5.3.2.1. CONSTRUCTION OF MICROARRAYS

Microarrays are prepared by selecting probes which comprise a polynucleotide sequence, and then immobilizing such probes to a solid support or surface. For example, the probes may comprise DNA sequences, RNA sequences, or copolymer sequences of DNA and RNA. The polynucleotide sequences of the probes may also comprise DNA and/or RNA analogues, or combinations thereof. For example, the polynucleotide sequences of the probes may be full or partial fragments of genomic DNA. The polynucleotide sequences of the probes may also be synthesized nucleotide sequences, such as synthetic oligonucleotide sequences. The probe sequences can be synthesized either enzymatically *in vivo*, enzymatically *in vitro* (e.g., by PCR), or non-enzymatically *in vitro*.

The probe or probes used in the methods of the invention are preferably immobilized to a solid support which may be either porous or non-porous. For example, the probes of the invention may be polynucleotide sequences which are attached to a nitrocellulose or nylon membrane or filter covalently at either the 3' or the 5' end of the polynucleotide. Such hybridization probes are well known in the art (see, e.g., Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). Alternatively, the solid support or surface may be a

glass or plastic surface. In a particularly preferred embodiment, hybridization levels are measured to microarrays of probes consisting of a solid phase on the surface of which are immobilized a population of polynucleotides, such as a population of DNA or DNA mimics, or, alternatively, a population of RNA or RNA mimics. The solid phase may be a nonporous or, optionally, a porous material such as a gel.

In preferred embodiments, a microarray comprises a support or surface with an ordered array of binding (e.g., hybridization) sites or "probes" each representing one of the markers described herein. Preferably the microarrays are addressable arrays, and more preferably positionally addressable arrays. More specifically, each probe of the array is preferably located at a known, predetermined position on the solid support such that the identity (i.e., the sequence) of each probe can be determined from its position in the array (i.e., on the support or surface). In preferred embodiments, each probe is covalently attached to the solid support at a single site.

Microarrays can be made in a number of ways, of which several are described below. However produced, microarrays share certain characteristics. The arrays are reproducible, allowing multiple copies of a given array to be produced and easily compared with each other. Preferably, microarrays are made from materials that are stable under binding (e.g., nucleic acid hybridization) conditions. The microarrays are preferably small, e.g., between 1 cm² and 25 cm², between 12 cm² and 13 cm², or 3 cm². However, larger arrays are also contemplated and may be preferable, e.g., for use in screening arrays. Preferably, a given binding site or unique set of binding sites in the microarray will specifically bind (e.g., hybridize) to the product of a single gene in a cell (e.g., to a specific mRNA, or to a specific cDNA derived therefrom). However, in general, other related or similar sequences will cross hybridize to a given binding site.

The microarrays of the present invention include one or more test probes, each of which has a polynucleotide sequence that is complementary to a subsequence of RNA or DNA to be detected. Preferably, the position of each probe on the solid surface is known. Indeed, the microarrays are preferably positionally addressable arrays. Specifically, each probe of the array is preferably located at a known, predetermined position on the solid support such that the identity (i.e., the sequence) of each probe can be determined from its position on the array (i.e., on the support or surface).

According to the invention, the microarray is an array (*i.e.*, a matrix) in which each position represents one of the markers described herein. For example, each position can contain a DNA or DNA analogue based on genomic DNA to which a particular RNA or cDNA transcribed from that genetic marker can specifically hybridize. The DNA or DNA analogue can be, *e.g.*, a synthetic oligomer or a gene fragment. In one embodiment, probes representing each of the markers is present on the array. In a preferred embodiment, the array comprises probes for each of the markers listed in Tables 1-5.

5.3.2.2. PREPARING PROBES FOR MICROARRAYS

As noted above, the "probe" to which a particular polynucleotide molecule specifically hybridizes according to the invention contains a complementary genomic polynucleotide sequence. The probes of the microarray preferably consist of nucleotide sequences of no more than 1,000 nucleotides. In some embodiments, the probes of the array consist of nucleotide sequences of 10 to 1,000 nucleotides. In a preferred embodiment, the nucleotide sequences of the probes are in the range of 10-200 nucleotides in length and are genomic sequences of a species of organism, such that a plurality of different probes is present, with sequences complementary and thus capable of hybridizing to the genome of such a species of organism, sequentially tiled across all or a portion of such genome. In other specific embodiments, the probes are in the range of 10-30 nucleotides in length, in the range of 10-40 nucleotides in length, in the range of 20-50 nucleotides in length, in the range of 40-80 nucleotides in length, in the range of 50-150 nucleotides in length, in the range of 80-120 nucleotides in length, and most preferably are 60 nucleotides in length.

The probes may comprise DNA or DNA "mimics" (*e.g.*, derivatives and analogues) corresponding to a portion of an organism's genome. In another embodiment, the probes of the microarray are complementary RNA or RNA mimics. DNA mimics are polymers composed of subunits capable of specific, Watson-Crick-like hybridization with DNA, or of specific hybridization with RNA. The nucleic acids can be modified at the base moiety, at the sugar moiety, or at the phosphate backbone. Exemplary DNA mimics include, *e.g.*, phosphorothioates.

DNA can be obtained, *e.g.*, by polymerase chain reaction (PCR) amplification of genomic DNA or cloned sequences. PCR primers are preferably chosen based on a known sequence of the genome that will result in amplification of specific fragments of genomic

DNA. Computer programs that are well known in the art are useful in the design of primers with the required specificity and optimal amplification properties, such as *Oligo* version 5.0 (National Biosciences). Typically each probe on the microarray will be between 10 bases and 50,000 bases, usually between 300 bases and 1,000 bases in length. PCR methods are well known in the art, and are described, for example, in Innis *et al.*, eds., PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS, Academic Press Inc., San Diego, CA (1990). It will be apparent to one skilled in the art that controlled robotic systems are useful for isolating and amplifying nucleic acids.

An alternative, preferred means for generating the polynucleotide probes of the microarray is by synthesis of synthetic polynucleotides or oligonucleotides, *e.g.*, using N-phosphonate or phosphoramidite chemistries (Froehler *et al.*, *Nucleic Acid Res.* 14:5399-5407 (1986); McBride *et al.*, *Tetrahedron Lett.* 24:246-248 (1983)). Synthetic sequences are typically between about 10 and about 500 bases in length, more typically between about 20 and about 100 bases, and most preferably between about 40 and about 70 bases in length. In some embodiments, synthetic nucleic acids include non-natural bases, such as, but by no means limited to, inosine. As noted above, nucleic acid analogues may be used as binding sites for hybridization. An example of a suitable nucleic acid analogue is peptide nucleic acid (*see, e.g.*, Egholm *et al.*, *Nature* 363:566-568 (1993); U.S. Patent No. 5,539,083).

Probes are preferably selected using an algorithm that takes into account binding energies, base composition, sequence complexity, cross-hybridization binding energies, and secondary structure. *See* Friend *et al.*, International Patent Publication WO 01/05935, published January 25, 2001; Hughes *et al.*, *Nat. Biotech.* 19:342-7 (2001).

A skilled artisan will also appreciate that positive control probes, *e.g.*, probes known to be complementary and hybridizable to sequences in the target polynucleotide molecules, and negative control probes, *e.g.*, probes known to not be complementary and hybridizable to sequences in the target polynucleotide molecules, should be included on the array. In one embodiment, positive controls are synthesized along the perimeter of the array. In another embodiment, positive controls are synthesized in diagonal stripes across the array. In still another embodiment, the reverse complement for each probe is synthesized next to the position of the probe to serve as a negative control. In yet another embodiment, sequences from other species of organism are used as negative controls or as "spike-in" controls.

5.3.2.3. ATTACHING PROBES TO THE SOLID SURFACE

The probes are attached to a solid support or surface, which may be made, *e.g.*, from glass, plastic (*e.g.*, polypropylene, nylon), polyacrylamide, nitrocellulose, gel, or other porous or nonporous material. A preferred method for attaching the nucleic acids to a surface is by printing on glass plates, as is described generally by Schena *et al.*, *Science* 270:467-470 (1995). This method is especially useful for preparing microarrays of cDNA (See also, DeRisi *et al.*, *Nature Genetics* 14:457-460 (1996); Shalon *et al.*, *Genome Res.* 6 :639-645 (1996); and Schena *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:10539-11286 (1995)).

A second preferred method for making microarrays is by making high-density oligonucleotide arrays. Techniques are known for producing arrays containing thousands of oligonucleotides complementary to defined sequences, at defined locations on a surface using photolithographic techniques for synthesis *in situ* (see, Fodor *et al.*, 1991, *Science* 251:767-773; Pease *et al.*, 1994, *Proc. Natl. Acad. Sci. U.S.A.* 91:5022-5026; Lockhart *et al.*, 1996, *Nature Biotechnology* 14:1675; U.S. Patent Nos. 5,578,832; 5,556,752; and 5,510,270) or other methods for rapid synthesis and deposition of defined oligonucleotides (Blanchard *et al.*, *Biosensors & Bioelectronics* 11:687-690). When these methods are used, oligonucleotides (*e.g.*, 60-mers) of known sequence are synthesized directly on a surface such as a derivatized glass slide. Usually, the array produced is redundant, with several oligonucleotide molecules per RNA.

Other methods for making microarrays, *e.g.*, by masking (Maskos and Southern, 1992, *Nuc. Acids. Res.* 20:1679-1684), may also be used. In principle, and as noted *supra*, any type of array, for example, dot blots on a nylon hybridization membrane (see Sambrook *et al.*, *MOLECULAR CLONING - A LABORATORY MANUAL* (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)) could be used. However, as will be recognized by those skilled in the art, very small arrays will frequently be preferred because hybridization volumes will be smaller.

In one embodiment, the arrays of the present invention are prepared by synthesizing polynucleotide probes on a support. In such an embodiment, polynucleotide probes are attached to the support covalently at either the 3' or the 5' end of the polynucleotide.

In a particularly preferred embodiment, microarrays of the invention are manufactured by means of an ink jet printing device for oligonucleotide synthesis, *e.g.*, using

the methods and systems described by Blanchard in U.S. Pat. No. 6,028,189; Blanchard et al., 1996, *Biosensors and Bioelectronics* 11:687-690; Blanchard, 1998, in *Synthetic DNA Arrays in Genetic Engineering*, Vol. 20, J.K. Setlow, Ed., Plenum Press, New York at pages 111-123. Specifically, the oligonucleotide probes in such microarrays are preferably synthesized in arrays, e.g., on a glass slide, by serially depositing individual nucleotide bases in "microdroplets" of a high surface tension solvent such as propylene carbonate. The microdroplets have small volumes (e.g., 100 pL or less, more preferably 50 pL or less) and are separated from each other on the microarray (e.g., by hydrophobic domains) to form circular surface tension wells which define the locations of the array elements (i.e., the different probes). Microarrays manufactured by this ink-jet method are typically of high density, preferably having a density of at least about 2,500 different probes per 1 cm². The polynucleotide probes are attached to the support covalently at either the 3' or the 5' end of the polynucleotide.

5.3.2.4. TARGET POLYNUCLEOTIDE MOLECULES

The polynucleotide molecules which may be analyzed by the present invention (the "target polynucleotide molecules") may be from any clinically relevant source, but are expressed RNA or a nucleic acid derived therefrom (e.g., cDNA or amplified RNA derived from cDNA that incorporates an RNA polymerase promoter), including naturally occurring nucleic acid molecules, as well as synthetic nucleic acid molecules. In one embodiment, the target polynucleotide molecules comprise RNA, including, but by no means limited to, total cellular RNA, poly(A)⁺ messenger RNA (mRNA) or fraction thereof, cytoplasmic mRNA, or RNA transcribed from cDNA (i.e., cRNA; see, e.g., Linsley & Schelter, U.S. Patent Application No. 09/411,074, filed October 4, 1999, or U.S. Patent Nos. 5,545,522, 5,891,636, or 5,716,785). Methods for preparing total and poly(A)⁺ RNA are well known in the art, and are described generally, e.g., in Sambrook *et al.*, *MOLECULAR CLONING - A LABORATORY MANUAL* (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). In one embodiment, RNA is extracted from cells of the various types of interest in this invention using guanidinium thiocyanate lysis followed by CsCl centrifugation (Chirgwin *et al.*, 1979, *Biochemistry* 18:5294-5299). In another embodiment, total RNA is extracted using a silica gel-based column, commercially available examples of which include RNeasy (Qiagen, Valencia, California) and StrataPrep (Stratagene, La Jolla, California). In an alternative embodiment, which is preferred for *S. cerevisiae*, RNA is extracted from cells

using phenol and chloroform, as described in Ausubel *et al.*, eds., 1989, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Vol. III, Green Publishing Associates, Inc., John Wiley & Sons, Inc., New York, at pp. 13.12.1-13.12.5). Poly(A)⁺ RNA can be selected, *e.g.*, by selection with oligo-dT cellulose or, alternatively, by oligo-dT primed reverse transcription of
5 total cellular RNA. In one embodiment, RNA can be fragmented by methods known in the art, *e.g.*, by incubation with ZnCl₂, to generate fragments of RNA. In another embodiment, the polynucleotide molecules analyzed by the invention comprise cDNA, or PCR products of amplified RNA or cDNA.

In one embodiment, total RNA, mRNA, or nucleic acids derived therefrom, is isolated
10 from a sample taken from a person afflicted with breast cancer. Target polynucleotide molecules that are poorly expressed in particular cells may be enriched using normalization techniques (Bonaldo *et al.*, 1996, *Genome Res.* 6:791-806).

As described above, the target polynucleotides are detectably labeled at one or more nucleotides. Any method known in the art may be used to detectably label the target
15 polynucleotides. Preferably, this labeling incorporates the label uniformly along the length of the RNA, and more preferably, the labeling is carried out at a high degree of efficiency. One embodiment for this labeling uses oligo-dT primed reverse transcription to incorporate the label; however, conventional methods of this method are biased toward generating 3' end fragments. Thus, in a preferred embodiment, random primers (*e.g.*, 9-mers) are used in
20 reverse transcription to uniformly incorporate labeled nucleotides over the full length of the target polynucleotides. Alternatively, random primers may be used in conjunction with PCR methods or T7 promoter-based *in vitro* transcription methods in order to amplify the target polynucleotides.

In a preferred embodiment, the detectable label is a luminescent label. For example,
25 fluorescent labels, bioluminescent labels, chemiluminescent labels, and colorimetric labels may be used in the present invention. In a highly preferred embodiment, the label is a fluorescent label, such as a fluorescein, a phosphor, a rhodamine, or a polymethine dye derivative. Examples of commercially available fluorescent labels include, for example, fluorescent phosphoramidites such as FluorePrime (Amersham Pharmacia, Piscataway, N.J.),
30 Fluoredit (Millipore, Bedford, Mass.), FAM (ABI, Foster City, Calif.), and Cy3 or Cy5 (Amersham Pharmacia, Piscataway, N.J.). In another embodiment, the detectable label is a radiolabeled nucleotide.

In a further preferred embodiment, target polynucleotide molecules from a patient sample are labeled differentially from target polynucleotide molecules of a standard. The standard can comprise target polynucleotide molecules from normal individuals (*i.e.*, those not afflicted with breast cancer). In a highly preferred embodiment, the standard comprises target polynucleotide molecules pooled from samples from normal individuals or tumor samples from individuals having sporadic-type breast tumors. In another embodiment, the target polynucleotide molecules are derived from the same individual, but are taken at different time points, and thus indicate the efficacy of a treatment by a change in expression of the markers, or lack thereof, during and after the course of treatment (*i.e.*, chemotherapy, radiation therapy or cryotherapy), wherein a change in the expression of the markers from a poor prognosis pattern to a good prognosis pattern indicates that the treatment is efficacious. In this embodiment, different timepoints are differentially labeled.

5.3.2.5. HYBRIDIZATION TO MICROARRAYS

Nucleic acid hybridization and wash conditions are chosen so that the target polynucleotide molecules specifically bind or specifically hybridize to the complementary polynucleotide sequences of the array, preferably to a specific array site, wherein its complementary DNA is located.

Arrays containing double-stranded probe DNA situated thereon are preferably subjected to denaturing conditions to render the DNA single-stranded prior to contacting with the target polynucleotide molecules. Arrays containing single-stranded probe DNA (*e.g.*, synthetic oligodeoxyribonucleic acids) may need to be denatured prior to contacting with the target polynucleotide molecules, *e.g.*, to remove hairpins or dimers which form due to self complementary sequences.

Optimal hybridization conditions will depend on the length (*e.g.*, oligomer versus polynucleotide greater than 200 bases) and type (*e.g.*, RNA, or DNA) of probe and target nucleic acids. One of skill in the art will appreciate that as the oligonucleotides become shorter, it may become necessary to adjust their length to achieve a relatively uniform melting temperature for satisfactory hybridization results. General parameters for specific (*i.e.*, stringent) hybridization conditions for nucleic acids are described in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989), and in Ausubel *et al.*, CURRENT

PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994). Typical hybridization conditions for the cDNA microarrays of Schena *et al.* are hybridization in 5 X SSC plus 0.2% SDS at 65°C for four hours, followed by washes at 25°C in low stringency wash buffer (1 X SSC plus 0.2% SDS), followed by 10 minutes at 25°C in higher stringency wash buffer (0.1 X SSC plus 0.2% SDS) (Schena *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:10614 (1993)). Useful hybridization conditions are also provided in, e.g., Tijssen, 1993, *HYBRIDIZATION WITH NUCLEIC ACID PROBES*, Elsevier Science Publishers B.V.; and Kricka, 1992, *NONISOTOPIC DNA PROBE TECHNIQUES*, Academic Press, San Diego, CA.

- 10 Particularly preferred hybridization conditions include hybridization at a temperature at or near the mean melting temperature of the probes (e.g., within 51°C, more preferably within 21°C) in 1 M NaCl, 50 mM MES buffer (pH 6.5), 0.5% sodium sarcosine and 30% formamide.

5.3.2.6. SIGNAL DETECTION AND DATA ANALYSIS

- 15 When fluorescently labeled probes are used, the fluorescence emissions at each site of a microarray may be, preferably, detected by scanning confocal laser microscopy. In one embodiment, a separate scan, using the appropriate excitation line, is carried out for each of the two fluorophores used. Alternatively, a laser may be used that allows simultaneous specimen illumination at wavelengths specific to the two fluorophores and emissions from the two fluorophores can be analyzed simultaneously (see Shalon *et al.*, 1996, "A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization," *Genome Research* 6:639-645, which is incorporated by reference in its entirety for all purposes). In a preferred embodiment, the arrays are scanned with a laser fluorescent scanner with a computer controlled X-Y stage and a microscope objective.
- 20 Sequential excitation of the two fluorophores is achieved with a multi-line, mixed gas laser and the emitted light is split by wavelength and detected with two photomultiplier tubes. Fluorescence laser scanning devices are described in Schena *et al.*, *Genome Res.* 6:639-645 (1996), and in other references cited herein. Alternatively, the fiber-optic bundle described by Ferguson *et al.*, *Nature Biotech.* 14:1681-1684 (1996), may be used to monitor mRNA
- 25 abundance levels at a large number of sites simultaneously.
- 30

Signals are recorded and, in a preferred embodiment, analyzed by computer, *e.g.*, using a 12 or 16 bit analog to digital board. In one embodiment the scanned image is despeckled using a graphics program (*e.g.*, Hijaak Graphics Suite) and then analyzed using an image gridding program that creates a spreadsheet of the average hybridization at each wavelength at each site. If necessary, an experimentally determined correction for "cross talk" (or overlap) between the channels for the two fluors may be made. For any particular hybridization site on the transcript array, a ratio of the emission of the two fluorophores can be calculated. The ratio is independent of the absolute expression level of the cognate gene, but is useful for genes whose expression is significantly modulated in association with the different breast cancer-related condition.

5.4. THERAPEUTIC REGIMENS SPECIFIC TO PATIENT SUBSETS

The benefit of identifying subsets of individuals that have a common condition, followed by identification of sets of genes informative for those particular subsets of individuals, is that such subdivision and identification tends to more accurately identify the subset of genes responsible for, or most closely associated with, a particular form of the condition. For example, breast cancer is a complex condition brought about by several different molecular mechanisms. ER+ individuals, particularly ER+, ER/AGE high individuals, show an increased level of expression of cell cycle-control genes, and the expression of these genes is highly informative for prognosis in this patient subset (*see* Examples). In ER⁻ individuals, however, the expression of these genes is not informative for prognosis.

The set of informative markers, therefore, can be used to assign a particular course of therapy to an individual, *e.g.*, an individual having breast cancer, depending upon the condition subset into which the individual is classified. In one embodiment, therefore, the invention provides a method of assigning a course of therapy to an individual having a condition, said method comprising classifying the individual into one of a plurality of subsets of a condition, wherein a plurality of informative genes has been identified for at least one of said subsets; and assigning a course of therapy known or suspected to be effective for treating the subset of the condition associated with those genes. In a specific embodiment, said condition is breast cancer, said patient subset is ER+, ER/AGE high status, and said course of therapy comprises the administration of one or more compounds known or suspected to be

effective at arresting the cell cycle. In a more specific embodiment, said one or more compounds comprises taxol or a vinca alkaloid.

Of course, any course of therapy selected or assigned on the basis of the above phenotypes and gene expression may be supplemented by other treatments or courses of therapy relevant to or known or suspected to be effective in the treatment of the condition. For example, the treatment of breast cancer may additionally comprise surgery, either tissue-preserving or radical, radiation treatment, chemotherapy other than that suggested by gene expression analysis, or any other therapy or treatment known or suspected to be effective.

5.5. CLINICAL TRIALS AND EPIDEMIOLOGICAL STUDIES

The method of the present invention may also be used to assign individuals to categories within a clinical trial, epidemiological study or the like. For example, individuals may be distinguished according to a characteristic of a condition, such as the presence or absence of specific proteins (e.g., estrogen receptor) or tissue structures (e.g., lymph nodes), and with prognosis, and the results of the trial correlated with prognosis. In a specific example, the condition is breast cancer, the characteristic is the presence of the estrogen receptor, and the prognosis is the expected occurrence or non-occurrence of metastases within a given period, for example, five years, after initial diagnosis. In another specific example, the condition is obesity, the characteristics are 24-hour energy expenditure, and the prognosis is the expected occurrence of heart disease or diabetes. In another specific example, the condition is a neurodegenerative disease, the characteristic is exposure to a particular range of concentration of an environmental toxin, and the prognosis is expected occurrence or degree of loss of motor function. In each case, the characteristics and expected outcome are used to assign the individual to a category within a clinical trial or epidemiological study.

Thus, the invention provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising classifying the individual into one of a plurality of condition categories differentiated by at least one genotypic or phenotypic characteristic of the condition; determining the level of expression, in a sample derived from said individual, of a plurality of genes informative for said condition category; determining whether said level of expression of said plurality of genes indicates that the individual has a good prognosis or a poor prognosis; and assigning the individual to a category in a clinical trial on the basis of prognosis.

In a specific embodiment, the invention provides a method of assigning an individual to a category in a breast cancer clinical trial, said method comprising: (a) classifying said individual as ER⁻, *BRCA1*, ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining for said individual the level of expression of at least two genes for which markers are listed in Table 1 if said individual is classified as ER⁻, *BRCA1*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻; (c) determining whether said individual has a pattern of expression of said at least two genes that correlates with a good prognosis or a poor prognosis; and (d) assigning said individual to at least one category in a clinical trial if said individual has a good prognosis, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis. In a more specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a). In another more specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of any other clinical, phenotypic or genotypic characteristic of breast cancer. In another more specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis of breast cancer, and determining from the expression of said second plurality of genes, in addition to said first plurality of genes, whether said individual has a good prognosis or a poor prognosis.

5.6. KITS

The present invention further provides for kits comprising the marker sets above. The components of the kits of the present invention are preferably contained in sealed containers. In a preferred embodiment, the kit comprises a microarray ready for hybridization to target polynucleotide molecules. In specific embodiments, the kit may comprise any of the microarrays described in detail in Section 5.3.2. Where proteins are the target molecules, the kit preferably comprises a plurality of antibodies for binding to specific condition-related proteins, and means for identifying such binding (*e.g.*, means for performing a sandwich assay, ELISA, RIA, *etc.*). Such antibodies may be provided, for example, individually or as part of an antibody array. The kit may additionally comprise software for the data analyses described above, as described in detail in Section 5.7. The kit preferably contains one or

more controls. Such a control may be an artificial population of marker-related or marker-derived polynucleotides suitable for hybridization to a microarray, wherein the markers are related to or relevant to the condition of interest (for example, breast cancer). The control may also, or alternatively, be a set of expression values stored on a computer disk or other storage medium.

The kits of the invention may be primarily diagnostic in nature; that is, they may assist a physician or researcher in determining a characteristic, for example, the prognosis, of a condition of interest, the likely response to a therapeutic regimen, the likely outcome of exposure to an environmental condition, such as toxin exposure, etc. The kits of the invention may also be used to classify individuals, for example, to place individuals into different groups in a clinical trial. The use of each kit is determined by the markers, microarrays, controls, etc. included.

5.7. COMPUTER-FACILITATED ANALYSIS

The analytic methods described in the previous sections can be implemented by use of the following computer systems and according to the following programs and methods. A computer system comprises internal components linked to external components. The internal components of a typical computer system include a processor element interconnected with a main memory. For example, the computer system can be based on an Intel 8086-, 80386-, 80486-, Pentium™, or Pentium™-based processor with preferably 32 MB or more of main memory. The computer system may also be a Macintosh or a Macintosh-based system, but may also be a minicomputer or mainframe.

The external components preferably include mass storage. This mass storage can be one or more hard disks (which are typically packaged together with the processor and memory). Such hard disks are preferably of 1 GB or greater storage capacity. Other external components include a user interface device, which can be a monitor, together with an inputting device, which can be a "mouse", or other graphic input devices, and/or a keyboard. A printing device can also be attached to the computer.

Typically, a computer system is also linked to network link, which can be part of an Ethernet link to other local computer systems, remote computer systems, or wide area communication networks, such as the Internet. This network link allows the computer system to share data and processing tasks with other computer systems.

Loaded into memory during operation of this system are several software components, which are both standard in the art and special to the instant invention. These software components collectively cause the computer system to function according to the methods of this invention. These software components are typically stored on the mass storage device. A software component comprises the operating system, which is responsible for managing computer system and its network interconnections. This operating system can be, for example, of the Microsoft Windows® family, such as Windows 3.1, Windows 95, Windows 98, Windows 2000, or Windows NT, or may be of the Macintosh OS family, or may be UNIX, a UNIX derivative such as LINUX, or an operating system specific to a minicomputer or mainframe. The software component represents common languages and functions conveniently present on this system to assist programs implementing the methods specific to this invention. Many high or low level computer languages can be used to program the analytic methods of this invention. Instructions can be interpreted during run-time or compiled. Preferred languages include C/C++, FORTRAN and JAVA. Most preferably, the methods of this invention are programmed in mathematical software packages that allow symbolic entry of equations and high-level specification of processing, including some or all of the algorithms to be used, thereby freeing a user of the need to procedurally program individual equations or algorithms. Such packages include Mathlab from Mathworks (Natick, MA), Mathematica® from Wolfram Research (Champaign, IL), or S-Plus® from Math Soft (Cambridge, MA). Specifically, the software component includes the analytic methods of the invention as programmed in a procedural language or symbolic package.

The software to be included with the kit comprises the data analysis methods of the invention as disclosed herein. In particular, the software may include mathematical routines for marker discovery, including the calculation of similarity values between clinical categories (e.g., prognosis) and marker expression. The software may also include mathematical routines for calculating the similarity between sample marker expression and template marker expression, using array-generated fluorescence data, to determine the clinical classification of a sample.

Additionally, the software may also include mathematical routines for determining the prognostic outcome, and the recommended therapeutic regimen, for an individual with a condition of interest. In the specific example of breast cancer, the mathematical routines

would determine the prognostic outcome and recommended therapeutic regimen for an individual having breast cancer. Such breast cancer-specific software would include instructions for the computer system's processor to receive data structures that include the level of expression of five or more of the marker genes listed in any of Tables 1-5 in a breast cancer tumor sample obtained from the breast cancer patient; the mean level of expression of the same genes in a control or template; and the breast cancer patient's clinical information, including age, lymph node status and ER status. The software may additionally include mathematical routines for transforming the hybridization data and for calculating the similarity between the expression profile of the marker genes in the patient's breast cancer tumor sample and a template. In a specific embodiment, the software includes mathematical routines for calculating a similarity metric, such as a coefficient of correlation, representing the similarity between the expression profile for the marker genes in the patient's breast cancer tumor sample and the template, and expressing the similarity as that similarity metric.

The software preferably would include decisional routines that integrate the patient's clinical and marker gene expression data, and recommend a course of therapy. In one embodiment, for example, the software causes the processor unit to receive expression data for prognosis-related genes in the patient's tumor sample, calculate a metric of similarity of these expression values to the values for the same genes in a template, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to a prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen. In a specific example, the software additionally causes the processor unit to receive data structures comprising clinical information about the breast cancer patient. In a more specific example, such clinical information includes the patient's age, estrogen receptor status, and lymph node status.

The software preferably causes the processor unit to receive data structures comprising relevant phenotypic and/or genotypic characteristics of the particular condition of interest, and/or of an individual having that condition, and classifies the individual into a condition subset according to those characteristics. The software then causes the processor to receive values for subset-specific markers, to calculate a metric of similarity of the values associated with those markers (*e.g.*, level, abundance, activity, *etc.*) from the individual to a template, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to a prognostic group, and,

on the basis of the prognostic group, assign a recommended therapeutic regimen. In the specific example of breast cancer and a breast cancer patient, the software, in one embodiment, causes the processor unit to receive data structures comprising the patient's age, estrogen receptor status, and lymph node status, and on the basis of this data, to classify the patient into one of the following patient subsets: ER⁻, sporadic; ER⁻, *BRCAl*; ER⁺, AR/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻. The software then causes the processor to receive expression values for subset-specific prognosis-informative gene expression in the patient's tumor sample, calculate a metric of similarity of these expression values to the values for the same genes in a patient subset-specific template or template, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to the prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen.

Where the template is an expression template comprising expression values for marker genes within a group of patients, e.g., breast cancer patients, the template can comprise either hybridization data obtained at the same time (i.e., in the same hybridization experiment) as the patient's individual hybridization data, or can be a set of hybridization or marker expression values stored on a computer, or on computer-readable media. If the latter is used, new patient hybridization data for the selected marker genes, obtained from initial or follow-up tumor samples, or suspected tumor samples, can be compared to the stored values for the same genes. However, the software may additionally comprise routines for updating the template data set, e.g., to add information from additional breast cancer patients or to remove existing members of the template data set, and, consequently, for recalculating the average expression level values that comprise the template. In another specific embodiment, said template comprises a set of single-channel mean hybridization intensity values for each of said at least five of said genes, stored on a computer-readable medium.

Clinical data relating to a breast cancer patient, or a patient having another type of condition, and used by the computer program products of the invention, can be contained in a database of clinical data in which information on each patient is maintained in a separate record, which record may contain any information relevant to the patient, the patient's medical history, treatment, prognosis, or participation in a clinical trial or study, including expression profile data generated as part of an initial diagnosis or for tracking the progress of the condition, for example, breast cancer, during treatment.

Thus, one embodiment of the invention provides a computer program product for classifying a breast cancer patient according to prognosis, the computer program product for use in conjunction with a computer having a memory and a processor, the computer program product comprising a computer readable storage medium having a computer program

5 mechanism encoded thereon, wherein said computer program product can be loaded into the one or more memory units of a computer and causes the one or more processor units of the computer to execute the steps of (a) receiving a first data structure comprising said breast cancer patient's age, ER status, LN status and tumor type; (b) classifying said patient as ER⁻, sporadic; ER⁻, *BRCAl*; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE

10 low, LN⁻; (c) receiving a second data structure comprising the levels of expression of at least two genes in a cell sample taken from said breast cancer patient wherein markers for said at least two genes are listed in Table 1 if said patient is classified as ER⁻, sporadic; Table 2 if said patient is classified as ER⁻, sporadic; Table 3 if said patient is classified as ER⁺, ER/AGE high; Table 4 if said patient is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if

15 said patient is classified as ER⁺, ER/AGE high, LN⁻; (d) determining the similarity of the expression profile of said at least two genes to a template expression profile of said at least two genes to obtain a patient similarity value; (e) comparing said patient similarity value to selected first and second threshold values of similarity, respectively, wherein said second similarity threshold indicates greater similarity to said template expression profile than does

20 said first similarity threshold; and (f) classifying said breast cancer patient as having a first prognosis if said patient similarity value exceeds said said second threshold similarity values, a second prognosis if said patient similarity value exceeds said first threshold similarity value but does not exceed said second threshold similarity value, and a third prognosis if said patient similarity value does not exceed said first threshold similarity value. In a specific

25 embodiment of said computer program product, said first threshold value of similarity and said second threshold value of similarity are values stored in said computer. In another more specific embodiment, said first prognosis is a "very good prognosis," said second prognosis is an "intermediate prognosis," and said third prognosis is a "poor prognosis," and wherein said computer program mechanism may be loaded into the memory and further cause said one or

30 more processor units of said computer to execute the step of assigning said breast cancer patient a therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and expression profile. In another specific embodiment, said computer program mechanism may

be loaded into the memory and further cause said one or more processor units of the computer to execute the steps of receiving a data structure comprising clinical data specific to said breast cancer patient. In a more specific embodiment, said single-channel hybridization intensity values are log transformed. The computer implementation of the method, however, may use any desired transformation method. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating the correlation between the expression profile of said genes in said cell sample taken from said breast cancer patient and the template expression profile of the same genes. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating the distance between the expression profile of said genes in said cell sample taken from said breast cancer patient and the template expression profile of the same genes.

Of course, the above breast cancer-specific examples are not limiting; analogous computer systems, software, and data analysis methods may be utilized for any condition of interest. For example, analogous software may be used to determine the prognosis of any other type of cancer, or of any other non-cancer diseases or conditions, using markers, expression level data and controls specific for that cancer, non-cancer disease or condition.

In an exemplary implementation, to practice the methods of the present invention, a user first loads experimental data into the computer system. These data can be directly entered by the user from a monitor, keyboard, or from other computer systems linked by a network connection, or on removable storage media such as a CD-ROM, floppy disk (not illustrated), tape drive (not illustrated), ZIP® drive (not illustrated) or through the network. Next the user causes execution of expression profile analysis software which performs the methods of the present invention.

In another exemplary implementation, a user first loads experimental data and/or databases into the computer system. This data is loaded into the memory from the storage media or from a remote computer, preferably from a dynamic geneset database system, through the network. Next the user causes execution of software that performs the steps of the present invention.

Additionally, because the data obtained and analyzed in the software and computer system products of the invention may be confidential, the software and/or computer system

preferably comprises access controls or access control routines, such as password protection and preferably, particularly if information is to be transmitted between computers, for example, over the Internet, encryption of the data by a suitable encryption algorithm (e.g., PGP).

Alternative computer systems and software for implementing the analytic methods of this invention will be apparent to one of skill in the art and are intended to be comprehended within the accompanying claims. In particular, the accompanying claims are intended to include the alternative program structures for implementing the methods of this invention that will be readily apparent to one of skill in the art.

6. EXAMPLES

6.1. EXAMPLE 1: IDENTIFICATION OF PHENOTYPIC SUBSETS AND INFORMATIVE GENESETS FOR EACH

Materials and Methods

Tumor Samples:

311 cohort samples were collected from breast cancer patients. Selection criteria for sporadic patients (*i.e.*, those not identified as having a *BRCA1*-type tumor; $n = 291$) included: primary invasive breast carcinoma less than 5 cm (T1 or T2); no axillary metastases (N0); age at diagnosis of less than 55 years; calendar year of diagnosis 1983-1996; and no previous malignancies. All patients were treated by modified radical mastectomy or breast-conserving treatment. See van't Veer *et al.*, *Nature* 415:530 (2002). Selection criteria for hereditary (*i.e.*, *BRCA1*-type; $n = 20$) tumors included: carriers of germline mutation in *BRCA1* or *BRCA2*, and primary invasive breast carcinoma. van't Veer, *supra*. Additionally, for development of a classifier for the *BRCA1* group, 14 *BRCA1* samples previously identified (see van't Veer, *supra*) were added to the 20 *BRCA1* type samples to increase sample size. Those 14 samples also satisfy the conditions that they are ER negative and age less than 55 years old.

Data analysis:

Sample sub-grouping: As shown in FIG. 1, tumor samples were first divided into ER+ and ER- branches since this is the dominant gene expression pattern. In the ER- branch, the samples were further divided into “BRCA1 mutation like” and “Sporadic like” categories

using the expression templates and 100 genes previously identified as optimal for determining *BRCA1* status. See van't Veer *et al.*, *Nature* 415:530 (2002). In the ER+ category, samples were divided by ER vs. age distribution (see below) into two groups, "ER/AGE low" and "ER/AGE high." Within the "ER/AGE low" group, samples were

5 further divided according to the lymph node status into two sub-groups: lymph node negative (0 lymph nodes; LN-) and positive (> 0 lymph nodes; LN+) group.

The result of these divisions was five distinctive sub-groups: "ER⁻, sporadic" (*n* = 52), "ER⁻, BRCA1" (*n* = 34), "ER+, ER/AGE high" (*n* = 83), "ER+, ER/AGE low, LN⁻" (*n* = 81), and "ER+, ER/AGE low, LN+" (*n* = 75). A few samples with a specific ER vs. age distribution

10 in "ER+, ER/AGE low, LN+" group were further excluded to develop a classifier, see below for details.

Estrogen receptor level: Estrogen receptor gene expression level was measured by a 60mer oligo-nucleotide on a microarray. Since every individual sample was compared to a pool of all samples, the ratio to pool was used to measure the relative level. A threshold of -0.65

15 on log₁₀(ratio) was used to separate the ER+ group from ER⁻ group. See van't Veer *et al.*, *Nature* 415:530 (2002).

Grouping by ER vs. age distribution: Samples were not uniformly distributed in ER vs. age space among the ER+ samples (FIG. 2). First, it appeared that the ER level increases with age, as there were few samples from young individuals having a high ER expression level.

20 For example, in the 35 to 40 years age group, samples having a log(ratio) of ER > 0.2 are relatively few as compared to the 40 to 45 age group. In the set of samples used, the 40 < age ≤ 45 group contains 30 samples having log(ratio) ER values between -0.2 to 0.2, and 28 samples having values greater than 0.2, whereas the 35 < age ≤ 40 group includes 24 samples with values between -0.2 to 0.2, but only 6 samples with values of greater than 0.2 (Fisher's

25 exact test P-value: 1%). The increasing ER level with age may simply due to the fact that estrogen levels decrease with age, and the estrogen receptor level rises in compensation.

There also appear to be at least two groups of patients, as indicated by the solid line separating the two in FIG. 2A. A bimodality test of the separation indicated by the solid line yielded P-value < 10⁻⁴. Each of these two groups has its own trend between the ER level and

30 age. The solid line can be approximated by ER = 0.1(age - 42.5). Patients having values

above the solid line are referred to as the "ER/AGE high" group, and the patients below the line as the "ER/AGE low" group.

Prognosis in each group:

Feature selection and performance evaluation: For the prognosis in each group, non-

- 5 informative genes were filtered in each group of patients. Specifically, only genes with $|\log_{10}(\text{ratio})| > \log_{10}(2)$ and P-value (for $\log(\text{ratio}) \neq 0$) < 0.01 in more than 3 experiments were kept. This step removed all genes that never had any significant change across all samples. The second step used a leave-one-out cross validation (LOOCV) procedure to optimize the number of reporter genes (features) in the classifier and to estimate the
- 10 performance of the classifier in each group. The feature selection was included inside the loop of each LOOCV process. The final "optimal" reporter genes were selected using all of the "training samples" as the result of "re-substitution" because one classifier was needed for each group.

- Selection of training samples: Only the samples from patients who had metastases within 5
- 15 years of initial diagnosis (3 years for "ER⁻, sporadic" samples; i.e., the "poor outcome" group), or who were metastases-free with more than 5 years of follow-up time (i.e., the "good outcome" group, were used as the training set. Because the average expression levels for informative genes among patients who were metastasis-free, or who had early metastases, were used as expression templates for prediction, the training samples for the ER⁺ samples
- 20 were further limited to those samples that could also be correctly classified by the first round of LOOCV process. For the "ER⁻, sporadic" samples, no such iteration was done because no improvement was observed. For the "ER⁻, BRCA1" samples, an iteration was done, but the training samples in the second iteration were limited to the correctly predicted good outcome samples from the first round of LOOCV, and all the poor outcome samples with
- 25 metastases time less than 5 years. Further limitation of the poor outcome samples was not performed because of the small number of poor samples and the absence of improvement by such limitation. In the first round of LOOCV, except for the "ER⁻, sporadic" group, the number of features was fixed at 50 genes. A patient was predicted to have a favorable outcome, that is, no metastases within five years of initial diagnosis, if the expression of the
- 30 reporter genes in a sample from the individual was more similar to the "average good profile" than the "average poor profile", and a poor outcome, that is, a metastasis within five years, if

the expression of the reporter genes in the sample was more similar to the “average poor profile” than the “average good profile”.

The justification for such an iteration operation is threefold. First, biologically, there are always a few individuals with specific reasons (different from the vast majority) to stay
15 metastases free or to develop metastases. Second, statistically, most groups of patients include outliers that don't follow the distribution of the majority of samples. Third, methodologically, the iteration operation is very similar to the idea of “boosting”, but instead of increasing the weights of the samples predicted wrong, emphasis is placed on the well behaved samples for selecting features and training the classifier. Since this process was
10 used to select “training samples”, and the performance was evaluated using the LOOCV (including the feature selection) after the training sample being fixed, there is no issue of over-fitting involved in our procedures. This method of iteration is thus more likely to reveal the dominant mode to metastases within each group.

Error rate and odds ratio, threshold in the final LOOCV: Unless otherwise stated, the error
15 rate is the average error rate from two populations: (1) the number of poor outcome samples misclassified as good outcome samples, divided by the total number of poor outcome samples; and (2) the total number of good outcome samples misclassified as poor outcome samples, divided by the total number of good samples. Two odds ratios are reported for a given threshold: (1) the overall odds ratio and (2) the 5 year odds ratio. The 5 year odds
20 ratio was calculated from samples from individuals that were metastases free for more than five years, and who experienced metastasis within 5 years. The threshold was applied to $\text{cor1} - \text{cor2}$, where “cor1” stands for the correlation to the “average good profile” in the training set, and “cor2” stands for the correlation to the “average poor profile” in the training set.

25 The threshold in the final round of LOOCV was defined using the following steps:
(1) For each of the N sample i left out for training, features based on the training set were selected. (2) Given a feature set, an incomplete LOOCV with N-1 samples was performed (only the “average poor profile” and “average good profile” is varied depending on whether the left out sample is in the training set or not). (3) The threshold based on the minimum
30 error rate from N-1 samples was determined, and that threshold was assigned to sample i in step (1). (4) The median threshold from all N samples was taken, and designated the final threshold. FIGS. 3-7 present detailed information about classifiers for the 5 groups: “ER⁻,”

sporadic", "ER⁻, BRCA1", "ER⁺, ER/age high", "ER⁺, ER/age low, LN⁻", "ER⁺, ER/age low, LN⁺". Tables 1-5 (see Section 5.3) list the final optimal reporter genes for each of the 5 classifiers for each of the five patient subsets. Table 6, below, summarizes the performance of each of the five classifiers together with thresholds used in each classifier.

5 Table 6. Performance of classifiers for each patient subset.

Classifier	Optimal # of Genes	(C1-C2) Threshold	Metastasis Free	# of Samples	TP	FP	FN	TN	Odds Ratio	95% C.I.
ER ⁺ , ER/AGE high	50	1.22	Overall	83	31	14	5	33	14.61	4.71-45.36
			5 year	71	24	11	3	33	24.00	6.03-95.46
ER ⁺ , ER/AGE low, LN ⁻	65	0.38	Overall	81	14	6	6	55	21.39	5.98-76.52
			5 year	73	11	4	5	53	29.15	6.73-126.33
ER ⁺ , ER/AGE low, LN ⁺	50	-0.12	Overall	56	7	4	6	39	11.38	2.54-50.94
			5 year	48	5	4	3	36	15.00	2.57-87.64
ER ⁻ , sporadic	20	-0.01	Overall	52	18	7	7	29	7.35	2.16-25.04
			5 year	45	16	5	6	18	9.60	2.45-37.58
ER ⁻ , BRCA1	10	-0.37	Overall	34	6	3	3	22	14.67	2.34-92.11
			5 year	22	6	1	3	12	24.00	2.04-282.68

TP: True positive

FP: False positive

FN: False negative

TN: True negative

- 10 Classification method: All classifiers described herein, feature selection and optimization were included inside the LOOCV loop. Classifier performance was based on the LOOCV results. The profile based on the selected features from each patient was compared to the "average good profile" and "average poor profile" (by correlation) to determine its predicted outcome.

- 15 Correlation calculation: The correlation between each gene's expression log(ratio) and the endpoint data (final outcome) was calculated using the Pearson's correlation coefficient. The correlation between each patient's profile and the "average good profile" and "average poor profile" is the cosine product (no mean subtraction).

Results:

- We employed the comprehensive prognosis strategy on microarray expression profiles of 311 patients diagnosed before age 55 that were all part of previous studies establishing and validating a 70-gene prognosis profile. *See van 't Veer et al., Nature* 415:530 (2002); van de Vijver *et al., N. Engl. J. Med.* 347:1999 (2002). In addition, 14 known *BRCA1* samples from the *Nature* study were included in defining the prognosis classifier for the *BRCA1* group. The overview of the stratifications is shown in FIG. 1. In each of the patient subsets, prognosis classifiers were developed and performance was evaluated by leave-one-out cross-validation. The biological make up of each of the classifiers was also examined.
- 10 During the process to decide whether a particular clinical parameter should be used for the next stratification, our objectives were twofold: (1) identification of homogeneous prognosis patterns; and/or (2) improved prognosis in the subsets. There is a subtle balance between these two objectives because smaller groups will likely lead to uniform patterns within the group but have increasingly limited predictive power. With the exception of the *BRCA1*
- 15 subset, each group in our stratification contained 50 or more samples.

- The first layer of stratification was based on the estrogen receptor level. We and others previously observed that estrogen receptor expression has a dominant effect on overall gene expression in breast cancer as seen in hierarchical clustering. *van 't Veer et al., Nature* 415:530 (2002); Perou *et al., Nature* 406:747 (2000); Gruvberger *et al., Cancer Res.* 61:5979
- 20 (2001). In our previous analysis up to 2500 genes are significantly correlated with ER expression levels in tumor. *van 't Veer et al., Nature* 415:530 (2002). According to the threshold defined previously (van de Vijver *et al., N. Engl. J. Med.* 347:1999 (2002)), samples were first divided into two groups according to the estrogen receptor level as measured by the oligo probe (accession number: NM_000125) on the array; samples with
- 25 $\log(\text{ratio}) > -0.65$ belong to the ER+ group, and the rest belong to ER- group). This resulted in 239 samples in the ER+ group and 72 samples in the ER- group.

- In the ER+ branch we observed that when displaying ER expression level as a function of age, at least two subgroups appear to exist. (In general, any bimodality in the clinical data is useful.) We therefore decided to stratify the tumors according this bimodality
- 30 (see FIG. 2). The group of ER+ patients having a high ER/AGE ratio was designated the "ER/AGE high" group (83 samples), and the remaining group of patients was designated "ER/AGE low" group (156 samples).

Within the "ER/age high" group, we identified a group of prognosis reporter genes that highly correlated with the outcome (*see* Table 3). Moreover, the expression of these genes appeared to be very homogeneous, as indicated by high similarity in expression among those genes. *See* FIG 2A. Leave-one-out cross validation including reporter selection
5 yielded an odds ratio of 14.6 (95%CI: 4.7-45.4) and 5 year odds ratio of 24.0 (95%CI: 6.0-95.5). Examination of those reporter genes reveals they are mostly the cell cycle genes which are highly expressed in the poor outcome tumors. It is worth noting that even though this group includes LN+ and LN- individuals, and mixed treatment, the incidence of distant metastases is predicted by a biologically uniform set of genes, possibly indicating that
10 proliferation is the prime driving force for disease progression. Also even though variation in these genes is observed in other tumor subgroups this is generally not correlated with outcome in those settings (*see* below).

In the "ER/age low" group, no predictive pattern was found in the whole group; thus, the samples were further stratified into LN- (81 samples, referred to as "ER/age low LN-")
15 and LN+ (75 samples, referred to as "ER/age low LN+") group.

Within the "ER/age low LN-" group, a group of genes was identified that was uniformly co-regulated, and which correlated with the outcome. Leave-one-out cross-validation (including feature selection) yielded an odds ratio of 21.4 (95% CI: 6.0-76.5) and 5 year odds ratio of 29.2 (95% CI: 6.7-126.3). This group of genes is also enriched for
20 individual biological functions (*see* below).

For the "ER/age low LN+" subset, an informative set of genes (*see* Table 4) was obtained after exclusion of several samples from older individuals having low ER levels. These samples are indicated in FIG. 2A as those lying below the dashed line (approximated as $ER < 0.1 * (age - 50)$). 56 samples remained after the exclusion. This sample set allowed the
25 identification of a group of genes with a highly homogeneous pattern that is useful for prognosis (overall odds ratio: 11.4 (2.5-50.9), 5 year odds ratio: 15.0 (2.6-87.6)). This suggests again that ER vs. age is an important combination for stratifying breast cancer patients. The reporter genes involved in this classifier also correlated with the clinical measure of the degree of lymphocytic infiltration (data not shown). The prediction in this
30 group is not as strong as other positive groups, which may indicate the primary tumor carries weaker information about the metastases for this group of patients, and the metastases may be started from or influenced by tumors already in lymph nodes.

In the ER⁻ branch, because a portion of the samples are “BRCA1-like,” it is natural to divide the samples into “BRCA1-like” and “sporadic like”. To perform the classification, the BRCA1/sporadic tumor type classifier described in Roberts *et al.*, “Diagnosis and Prognosis of Breast Cancer Patients,” International Publication No. WO 02/103320, which is hereby
5 incorporated by reference in its entirety, to segregate the ER⁻ cohort samples. 52 out of the 72 ER⁻ samples were found to be “sporadic like” and 20 were found to be “BRCA1-like”. Interestingly, the “sporadic like” group is enriched for *erbb2* mutations (data not shown).

Within the “ER⁻, sporadic” group, no homogeneous prognosis pattern was identified; however, 20 genes were identified that are highly predictive of the tumor outcome (see Table
10 2). Leave-one-out cross-validation including feature selection yielded an odds ratio of 7.4 (95% CI 2.2-25.0) and 5 year odds ratio 9.6 (2.5 – 37.6). This result represents a significant improvement in prognosis compared to the previously-identified 70 gene prognosis classifier (see Roberts *et al.*, International Publication No. WO 02/103320; van’t Veer *et al.*, *Nature* 415:530 (2002)) which has no within-group prognostic power for the ER⁻ patient subset.
15 The fact that 20 genes predict outcome and that there is no homogeneous (and apparent biological) pattern in this group probably indicates multiple mechanisms of metastasis in this group. Gene annotation indicates that genes included may be involved in invasion, energy metabolism and other functions.

For the “ER⁻, BRCA1-like” group, we added 14 BRCA1 mutation carrier samples
20 from our previous study to increase the number of samples. Those 14 extra samples also satisfy our selection criteria: ER negative and age less than 55 years. The leave-one-out cross validation process identified 10 genes that are predictive of final outcomes. The overall odds ratio is 14.7 (95% CI: 2.3-92.1) and the 5 year odds ratio is 24.0 (95% CI: 2.0-282.7).

Because no homogeneous gene expression patterns were found in ER⁻ branch, the
25 predictive power of those genes was further validated. One means of further validation was to review the different classifier gene sets for biological interpretations and to identify genes within each classifier that gave indications as to the origins of the tumors.

The “ER+, ER/AGE high” group yielded a classifier highly enriched for cell cycle genes with both G1/S and G2/M phases represented. In this group, over-expression of 46 of
30 the 50 genes is associated with disease progression including all the known cell cycle genes. This is consistent with rapid growth being the determinant of metastatic potential. Four

genes in this classifier are anti-correlated with outcome and cell cycle. One of these genes encodes follistatin, which binds to and inhibits activin and other members of the TGF β family (Lin *et al.*, *Reproduction* 126:133 (2003)), the members of which have many functions, including growth stimulation. Tumor grade also accurately predicts metastatic potential in this group (overall odds ratio: 5.9, 95% CI: 2.0-18.0, 5 year odds ratio: 12.5, 95% CI: 2.6-59.3) and is also correlated with the expression level of these genes, which is consistent with rate of growth being the primary determinant of disease progression. This set of genes has a significantly lower correlation with outcome in the other patient subsets, even though coordinate and similarly variable expression is seen. For example, many tumors in the "ER⁻, sporadic" group have high cell cycle and low FST expression, but the expression of these genes in these groups is minimally correlated with outcome, indicating that growth is not the primary determinant of outcome here (*see* FIGS. 8A and 8B).

The ER⁺, ER/AGE low, LN⁻ group yielded a classifier rich in both genes for glycolytic enzymes (12 of 56) and genes induced by hypoxia and/or angiogenesis (14 of 56) with 5 genes falling into both categories. These genes are positively correlated with poor outcome, implying that energy metabolism (glycolysis), angiogenesis and adaptation to hypoxia are critical pathways in this subgroup of tumors. None of these genes appear in the classifiers for the other patient subsets, and there is a much reduced predictive value of these genes in the other tumors, even though coordinate and similarly variable expression is seen (*see* FIG. 8C and 8D).

The implication of the above analyses is that certain well known functions (growth, angiogenesis, energy metabolism) are important in certain tumor types and not in others, and therefore therapies that target these functions will be likely be similarly effective in some tumor subgroups and not in others. For example therapies that target cell cycle progression, such as taxol or the vinca alkaloids, may be optimally effective in the ER⁺, ER/AGE high group, where overexpression of cell cycle genes predominates in the classifier. In contrast, tumor subgroups in which variation in cell cycle expression is not correlated with outcome may be less sensitive to taxol or the vinca alkaloids.

The "comprehensive prognosis" approach significantly improves the prediction error rate when compared with 70 gene classifier (Table 7). To make the comparison fair, we listed two sets of results from the 70 gene classifier. The first results from the use of the same threshold applied to all the patient subsets (threshold previously optimized for false

negative rate); the second one results from the use of a threshold optimized for each patient subset (optimized for average error rate). The comprehensive approach lowered the error rate by at least 6%.

Table 7. Average error rate for the patient subset approach compared with the previously-described 70 gene classifier.

Prognosis method	over all error rate	5 year error rate
70 gene, fix thresh	30.90%	25.70%
70 gene, opt thresh	28.60%	27.60%
Comprehensive	21.50%	19.30%

Fix thresh: use of a fixed threshold in the classifier as previously determined.

Opt threshold: use of a threshold optimized for each sub-group. For the "ER/Age low, LN+" subgroup, 56 samples used for developing the classifier were included here, resulted in 306 samples in total.

6.2. EXAMPLE 2: IDENTIFICATION OF CHEMOTHERAPY RESPONSE SUBPOPULATION IN BREAST CANCER PATIENTS

Predicting responders to a treatment is extremely important in daily clinical therapies, in order to select the right medicine. It is also crucial for pharmaceutical companies running clinical trials on new medicines to target the right population for efficacy. This Example shows that, by combining clinical parameters with gene expression prediction, a subpopulation of patients with characteristics of "ER/age low" and a "poor signature" has an enhanced response to chemotherapy treatment. On the other hand, patients with a poor signature but do not respond to chemotherapy are potential target for clinical trials of new anti-cancer medicines.

Introduction

Gene expression can be used to predict breast cancer outcome. A "70 gene" classifier (see Table 8) that can be used to identify patients with good outcome for avoiding over-treatment of chemotherapy was demonstrated (van't Veer LJ, Dai H, van de Vijver MJ, et al., Gene expression profiling predicts clinical outcome of breast cancer, Nature 2002; 415: 530-536). This classifier has been verified by a large cohort of samples (van de Vijver MJ, He YD, van't Veer LJ, et al., A gene-expression signature as a predictor of survival in breast

cancer, N Engl J Med 2002; 347: 1999-2009). The result also implies that the treatment effect will be more significant within the "predicted poor" group since patients in the "predicted good" group are already very good, and may not improve further by additional treatment.

- 5 Recently, it has been demonstrated that ER+ patients can be further divided into subpopulations. The cell proliferation signature has a very strong prognostic power in a group of patients characterized by a high estrogen receptor levels relative to their ages (termed "ER/age high" group), but reduced predictive power in the "ER/age low" group and no predictive power in the ER- patients (U.S. Provisional Patent Application No. 60/550,810
- 10 by Dai et al., filed March 5, 2004, attorney docket no. 9301-229-888, which is incorporated by reference in its entirety).

Data set

- A cohort of 311 breast cancer samples were used (van 't Veer LJ, Dai H, van de Vijver MJ, et al., Gene expression profiling predicts clinical outcome of breast cancer, Nature
- 15 2002; 415: 530-536; van de Vijver MJ, He YD, van't Veer LJ, et al., A gene-expression signature as a predictor of survival in breast cancer, N Engl J Med 2002; 347: 1999-2009). Those samples were from patients younger than 55 years of age at the time of cancer diagnosis. Among them, 110 patients received chemotherapy.

Results

- 20 The following section shows that by combining the microarray prediction with the patient stratification based on clinical parameters, a subpopulation of breast cancer patients with enhanced chemotherapy treatment response can be identified.

- The treatment response was measured by the improvement in the metastasis free probability at 10 years after the diagnosis, between the treated patients and the untreated
- 25 patients.

Chemotherapy has limited effect in breast cancer patients less than 55 years of age. Shown in Figure 9 is the metastasis free probability as a function of time after diagnosis (Kaplan-Meier plot) for the entire set of patients. Patients were divided into two groups according to whether or not they received chemotherapy. Within the entire set of 311

patients younger than 55 years, 110 received chemotherapy. The treatment effect is very small, with the difference in 10 year metastasis free probability being only ~3%.

The “70 gene” prognostic classifier enhances the treatment effect. The purpose of the “70 gene” prognostic classifier (Table 8) was to save patients from unnecessary chemotherapy if they have a “good signature”. In other words, there may not be further benefit to patients in a good prognosis patient group if they potentially have a very good outcome. Hence, treatment should be given to the “predicted poor” group since they could potentially benefit from it. Results in Figure 10 validate the idea. Figure 10a displays the treatment effect of the “predicted good” group. The group was selected by high correlation to the “average expression of good outcome patients” as determined by 70 gene marker set (column C1 of Table 1) with correlation > 0.5 . There were 84 patients within this group, and 44 received chemotherapy. The treatment showed no positive effect in reducing the metastasis rate.

However, the patients with “poor signature” demonstrate an improved treatment response (Figure 10b) as compared to Figure 9. There are 226 patients with “poor signature” (correlation < 0.5), and 81 received chemotherapy. The 10 year metastasis free probability for those received chemotherapy is 9% higher than those did not.

In this example, a threshold of 0.5 in correlation was used to select patients with “poor signature”. This threshold is higher than the threshold of 0.4 that was used previously (van't Veer LJ, Dai H, van de Vijver MJ, et al., Gene expression profiling predicts clinical outcome of breast cancer, *Nature* 2002; 415: 530-536; van de Vijver MJ, He YD, van't Veer LJ, et al., A gene-expression signature as a predictor of survival in breast cancer, *N Engl J Med* 2002; 347: 1999-2009), for the purpose of keeping as many patients in the “predicted poor” group as possible for further stratification. The results were very similar if the original threshold of 0.4 was used.

Limiting to ER+ patients gains extra treatment response. Breast cancer patients with positive estrogen receptor status (ER+) have a distinctively different gene expression pattern than those with negative status (ER-). The “70 gene” classifier has a good prognostic power in the ER+ patients, but little prognostic power in the ER- patients (almost all the ER- patients were predicted to have a poor outcome).

The treatment effect is also very different for ER- patients than for ER+ patients with “poor signature”. As shown in Figure 11a, there is no difference in metastasis rate between the treated and untreated patients in the ER- group.

The treatment effect for ER+ patients with a “poor signature” (Figure 11b) is further increased when compared to Figure 10b. There are 155 patients in this group, and 60 received chemotherapy. The metastasis free probability at 10 years for patients with treatment is 14% higher than those without.

“ER/age high” group is ideal for chemotherapy treatment. Recently, it was discovered that the ER+ patients can be further divided into two groups according to their ER levels relative to the age at diagnosis. As shown in Figure 12, patients with relatively high ER level to their age are termed “ER/AGE high” group, and patients with relatively low ER level to their age are termed as “ER/AGE low” group. The “ER/AGE high” group has a relatively poor outcome compared to the “ER/AGE low” group, but the outcome in the “ER/AGE high” group can be accurately predicted by the cell proliferation genes. The same set of genes has a reduced prognostic power in the “ER/AGE low” group.

As shown in Figure 13, chemotherapy has little effect in the patients of “ER/AGE high” group with a “poor signature”. Excluding this group resulted in a further improved effect in the “ER/AGE low” group. In the “ER/AGE low” group, there are 101 patients with “poor signature”, and 39 received chemotherapy. The difference in metastasis free probability at 10 years between the treated and untreated group is 20%.

In summary, by progressively excluding “non-responsive” patient groups as defined by prognosis signature and clinical parameters (ER and age), it is possible to improve the treatment effect (as reported by the increased metastasis free probability at 10 years after diagnosis) from 3% to 20% (Table 9). Roughly, for patients younger than 55 years, there are about 1/3 of patients that do not need a treatment (“good signature”), about 1/3 patients do not respond to the chemotherapy (ER- and “ER/AGE high”) and about 1/3 patients are “responders” that should be treated with chemotherapy (“ER/AGE low”).

Table 9: Treatment effect in various patient groups

	All patients	Patients with “poor signature”	ER+ & “poor signature”	ER+, ER/AGE high & “poor signature”
--	--------------	--------------------------------------	---------------------------	---

Difference in 10 year metastasis free probability with/without chemo	3%	9%	14%	20%
--	----	----	-----	-----

7. REFERENCES CITED

All references cited herein are incorporated herein by reference in their entirety and for all purposes to the same extent as if each individual publication or patent or patent
5 application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes.

Many modifications and variations of the present invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific
embodiments described herein are offered by way of example only, and the invention is to be
10 limited only by the terms of the appended claims along with the full scope of equivalents to which such claims are entitled.

WHAT IS CLAIMED IS:

1. A method for predicting responsiveness of a breast cancer patient to chemotherapy based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising predicting said patient to exhibit

(a) enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is low; or

(b) reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if (i) said ER level is ER⁻, or (ii) said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high;

wherein said cellular constituent profile is classified as a poor prognosis profile if said cellular constituent profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of good outcome patients and said poor prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-occurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a patient who has occurrence of metastases within a second period of time after initial diagnosis.

2. The method of claim 1, said method comprising predicting said patient to exhibit enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is low.

3. The method of claim 1, said method comprising predicting said patient to exhibit reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if said ER level is ER⁻.

5 4. The method of claim 1, said method comprising predicting said patient to exhibit reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high.

5. The method of claim 1, further comprising determining said cellular constituent profile, said ER level, and/or, said ER/AGE.

10 6. The method of claim 5, wherein said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality
15 of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

7. The method of claim 6, wherein said cellular constituent profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a
20 control sample.

8. The method of claim 1, wherein said cellular constituent profile comprises measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and
25 wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome patients.

9. The method of claim 7, wherein measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of
30 good outcome patients.

10. The method of claim 9, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a correlation coefficient between said cellular constituent profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold indicates a high similarity and said correlation coefficient
5 equal to or less than said correlation threshold indicates a low similarity.

11. The method of claim 9, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a distance between said cellular constituent profile and said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance equal to or greater than said given value indicates
10 a low similarity.

12. The method of claim 10, wherein said correlation threshold is 0.5.

13. The method of claim 12, wherein said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene in said control sample, and wherein said ER level is classified
15 as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and wherein said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65.

14. The method of claim 13, wherein said gene encoding said estrogen receptor is the estrogen receptor α gene.

15. The method of claim 14, wherein said ER/AGE is classified as high if said ER
20 level as measured by $\log_{10}(\text{ratio})$ is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

16. The method of claim 15, wherein said estrogen receptor level is measured by an oligonucleotide probe that detects a transcript corresponding to the gene having accession
25 number NM_000125, wherein said control sample is a pool of breast cancer cells of different patients, and wherein $c = 0.1$ and $d = 42.5$.

17. The method of claim 16, wherein said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients.

18. The method of claim 16, wherein said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript of said gene encoding said estrogen receptor.

19. The method of claim 17, wherein said plurality of transcripts consists of
5 transcripts corresponding to at least a portion of the set of genes listed in Table 8.

20. The method of claim 19, wherein said plurality of transcripts consists of all transcripts corresponding to genes listed in Table 8.

21. The method of claim 20, wherein said differential measurements is selected from the group consisting of xdev, log(ratio), error-weighted log(ratio), and mean subtracted
10 log(intensity).

22. The method of claim 21, wherein said first period is 10 years and said second period is 10 years.

23. The method of claim 22, wherein said patient is under the age of 55, and wherein each said good outcome patient and each said poor outcome patient are under the age of 55 at
15 time of diagnosis of breast cancer.

24. A method for selecting a patient for enrollment in a clinical trial of a drug for treating breast cancer based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or
20 estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising selecting a patient for inclusion in said clinical trial if (a) said ER level is ER⁺; or (b) said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high;

25 wherein said cellular constituent profile is classified as a poor prognosis profile if said cellular constituent profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of good outcome patients and said poor prognosis template

comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-occurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a breast cancer patient who has
5 occurrence of metastases within a second period of time after initial diagnosis.

25. The method of claim 24, said method comprising selecting a patient for inclusion in said clinical trial if said ER level is ER⁻.

26. The method of claim 24, said method comprising selecting a patient for inclusion in said clinical trial if said cellular constituent profile is a poor prognosis profile, said ER
10 level is ER⁺, and said ER/AGE is high.

27. The method of claim 24, further comprising determining said cellular constituent profile, said ER level, and/or, said ER/AGE.

28. The method of claim 27, wherein said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient, wherein said
15 good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

20 29. The method of claim 28, wherein said cellular constituent profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample.

30. The method of claim 24, wherein said cellular constituent profile comprises
25 measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome
30 patients.

31. The method of claim 29, wherein measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

32. The method of claim 31, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a correlation coefficient between said cellular constituent profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

33. The method of claim 31, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a distance between said cellular constituent profile and said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance greater than said given value indicates a low similarity.

34. The method of claim 32, wherein said correlation threshold is 0.5.

35. The method of claim 34, wherein said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene in said control sample, and wherein said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and wherein said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65.

36. The method of claim 35, wherein said gene encoding said estrogen receptor is the estrogen receptor α gene.

37. The method of claim 36, wherein said ER/AGE is classified as high if said ER level is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

38. The method of claim 37, wherein said estrogen receptor level is measured by an oligonucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, wherein said control sample is a pool of breast cancer cells, and wherein $c = 0.1$ and $d = 42.5$.

39. The method of claim 38, wherein said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients.

40. The method of claim 38, wherein said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript corresponding to said gene encoding said estrogen receptor.

41. The method of claim 39, wherein said plurality of transcripts consists of transcripts corresponding to at least a portion of the set of genes listed in Table 8.

42. The method of claim 41, wherein said plurality of genes consists of all transcripts corresponding to genes listed in Table 8.

10 43. The method of claim 42, wherein said differential measurements is selected from the group consisting of xdev, log(ratio), error-weighted log(ratio), and mean subtracted log(intensity).

44. The method of claim 43, wherein said first period is 10 years and said second period is 10 years.

15 45. The method of claim 44, wherein said patient is under the age of 55, and wherein each said good outcome patient and each poor outcome patient are under the age of 55 at time of diagnosis of breast cancer.

46. A method for identifying a breast cancer patient as a good or poor candidate for chemotherapy based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising

25 (a) determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, said breast cancer patient is identified as a good candidate for chemotherapy; or

(b) determining whether said ER level is ER⁻, wherein if said ER level is ER⁻, said the breast cancer patient is identified as a poor candidate for chemotherapy; or

- (c) determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, said breast cancer patient is identified as a poor candidate for chemotherapy.

47. The method of claim 46, said method comprising determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, said breast cancer patient is identified as a good candidate for chemotherapy.

48. The method of claim 46, said method comprising determining whether said ER level is ER⁻, wherein if said ER level is ER⁻, said the breast cancer patient is identified as a poor candidate for chemotherapy.

49. The method of claims 46, said method comprising determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, said breast cancer patient is identified as a poor candidate for chemotherapy.

50. The method of claim 46, further comprising determining said cellular constituent profile, said ER level, and/or, said ER/AGE.

51. The method of claim 50, wherein said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

52. The method of claim 51, wherein said cellular constituent profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample.

5 53. The method of claim 50, wherein said cellular constituent profile comprises measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein
10 species representative of levels of said protein species in said plurality of poor outcome patients.

54. The method of claim 52, wherein measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

15 55. The method of claim 54, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a correlation coefficient between said cellular constituent profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

20 56. The method of claim 55, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a distance between said cellular constituent profile and said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance greater than said given value indicates a low similarity.

25 57. The method of claim 56, wherein said correlation threshold is 0.5.

58. The method of claim 57, wherein said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene in said control sample, and wherein said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and wherein said ER level
30 is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65.

59. The method of claim 58, wherein said gene encoding said estrogen receptor is the estrogen receptor α gene.

60. The method of claim 59, wherein said ER/AGE is classified as high if said ER level as measured by $\log_{10}(\text{ratio})$ is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

61. The method of claim 60, wherein said estrogen receptor level is measured by an oligonucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, wherein said control sample is a pool of breast cancer cells, and wherein $c = 0.1$ and $d = 42.5$.

62. The method of claim 61, wherein said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients.

63. The method of claim 61, wherein said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript corresponding to said gene encoding said estrogen receptor.

64. The method of claim 61, wherein said plurality of transcripts consists of transcripts corresponding to at least a portion of the set of genes listed in Table 8.

65. The method of claim 64, wherein said plurality of genes consists of all transcripts corresponding to genes listed in Table 8.

66. The method of claim 65, wherein said differential measurements is selected from the group consisting of x_{dev} , $\log(\text{ratio})$, error-weighted $\log(\text{ratio})$, and mean subtracted $\log(\text{intensity})$.

67. The method of claim 66, wherein said first period is 10 years and said second period is 10 years.

68. The method of claim 67, wherein said patient is under the age of 55, and wherein each said good outcome patient and each poor outcome patient are under the age of 55 at time of diagnosis of breast cancer.

69. The method of any one of claims 1-68, wherein said chemotherapy is carried out using the CMF combination comprising cyclophosphamide, methotrexate, and 5-fluorouracil.

70. A computer system comprising

a processor, and

5 a memory coupled to said processor and encoding one or more programs,

wherein said one or more programs cause the processor to carry out the method of any one of claims 1-68.

71. A computer program product for use in conjunction with a computer having a processor and a memory connected to the processor, said computer program product
10 comprising a computer readable storage medium having a computer program mechanism encoded thereon, wherein said computer program mechanism may be loaded into the memory of said computer and cause said computer to carry out the method of any one of claims 1-68.

72. A method for predicting responsiveness of a breast cancer patient to chemotherapy based on one or more of the following (i) a cellular constituent profile
15 comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising predicting
20 said patient to exhibit

(a) enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is low; or

(b) reduced response to chemotherapy as compared to patients in the general
25 population of breast cancer patients if (i) said ER level is ER⁻, or (ii) said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high. wherein said cellular constituent profile is classified as a good prognosis profile if said cellular constituent profile predicts non-occurrence of metastases in said breast cancer patient within a predetermined period of time after initial diagnosis, and wherein said cellular

constituent profile is classified as a poor prognosis profile if said cellular constituent profile predicts occurrence of metastasis within said predetermined period of time.

1/15

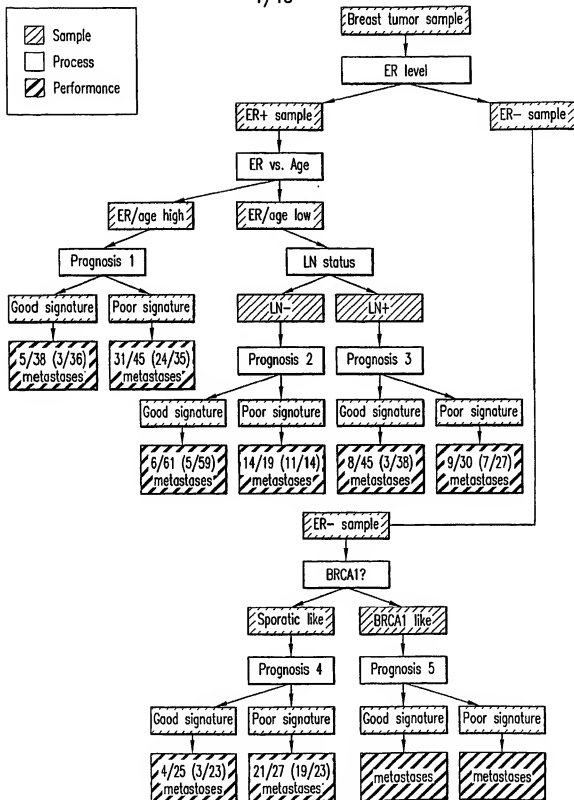


FIG. 1

2/15

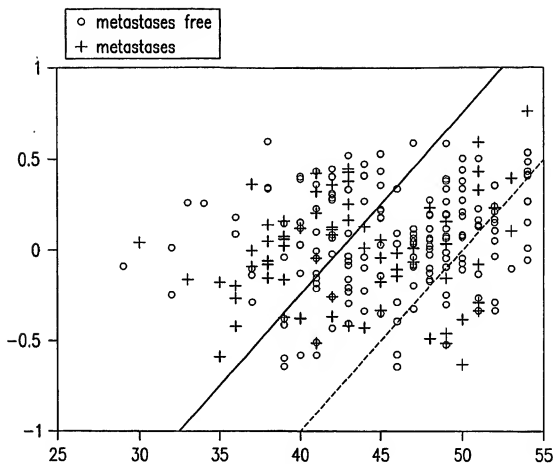


FIG.2A

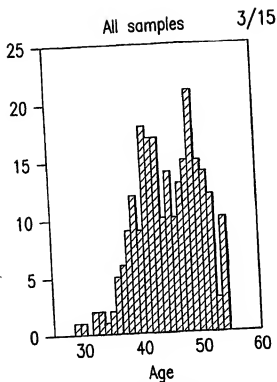


FIG.2B

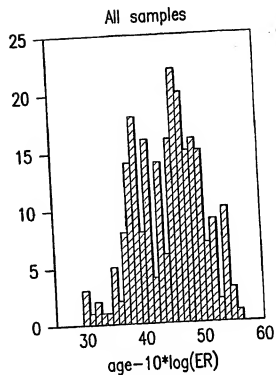


FIG.2C

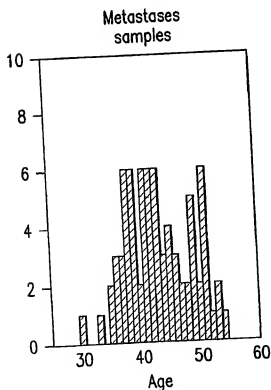


FIG.2D

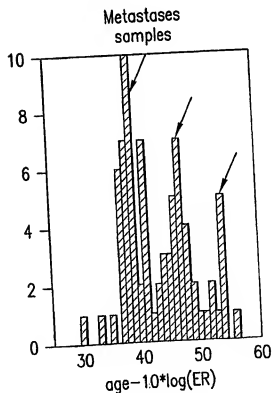


FIG.2E

4/15

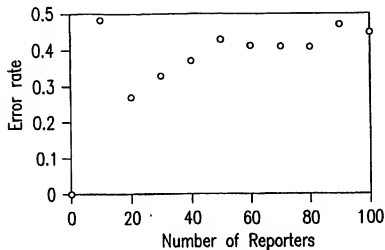


FIG.3A

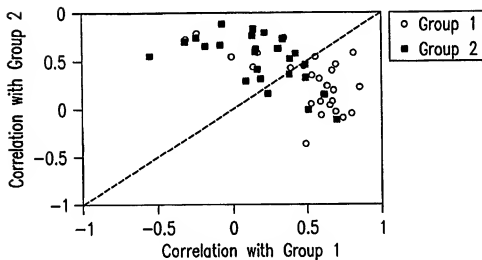


FIG.3B

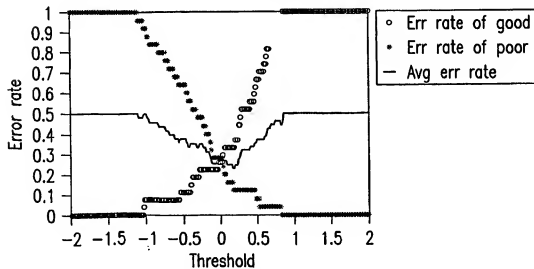


FIG.3C

5/15

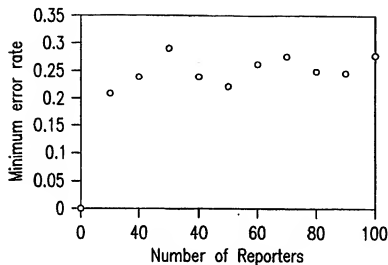


FIG.4A

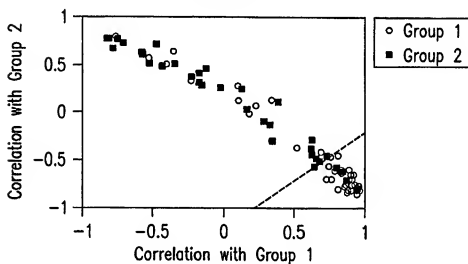
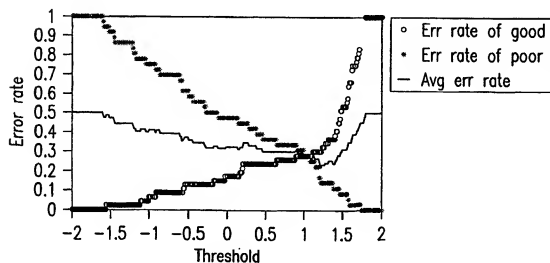


FIG.4B

FIG.4C
SUBSTITUTE SHEET (RULE 26)

6/15

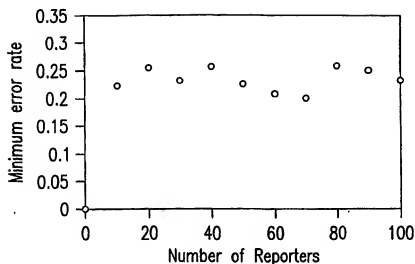


FIG.5A

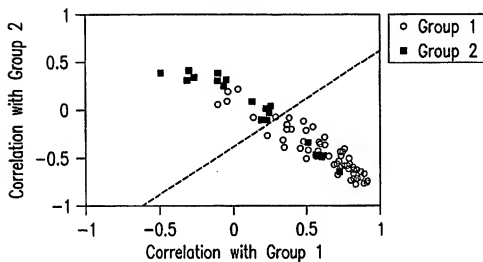
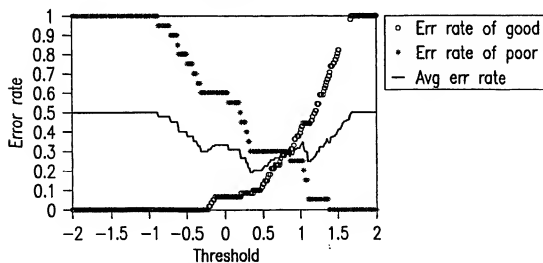


FIG.5B

FIG.5C
SUBSTITUTE SHEET (RULE 26)

7/15

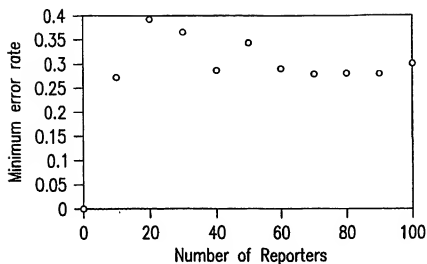


FIG.6A

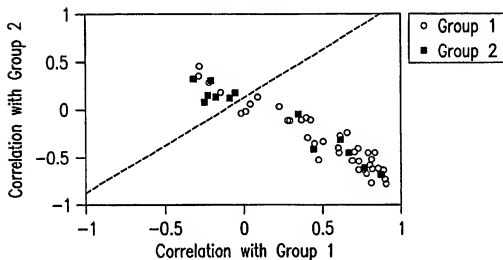


FIG.6B

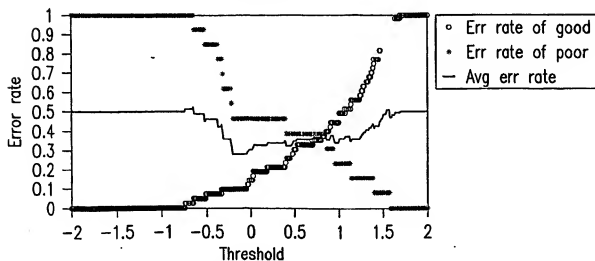


FIG.6C

8/15

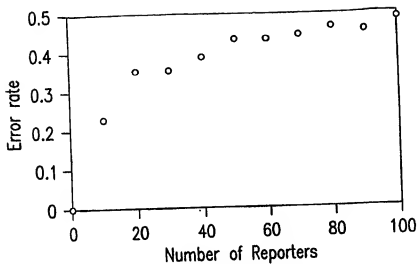


FIG.7A

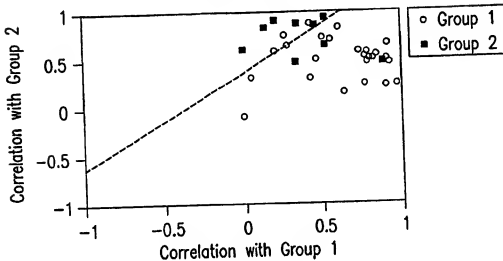
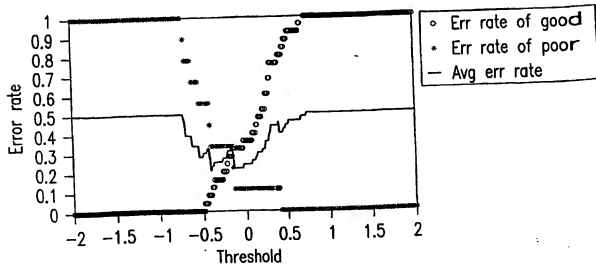


FIG.7B

FIG.7C
SUBSTITUTE SHEET (RULE 26)

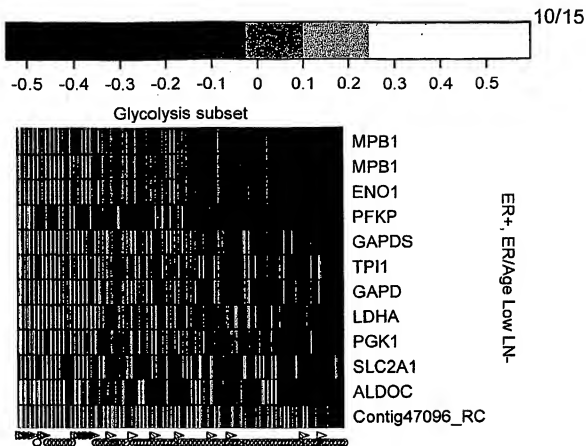
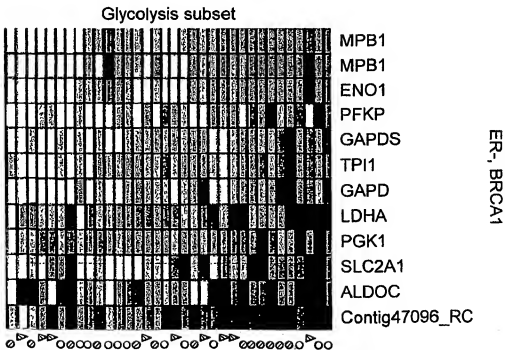


FIG. 8C

FIG. 8D
SUBSTITUTE SHEET (RULE 26)

11/15

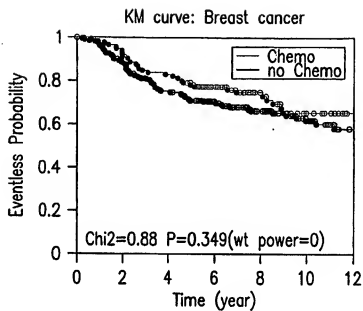


FIG.9

12/15

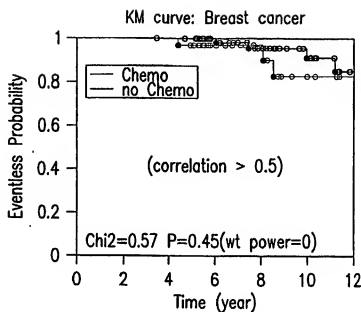


FIG. 10A

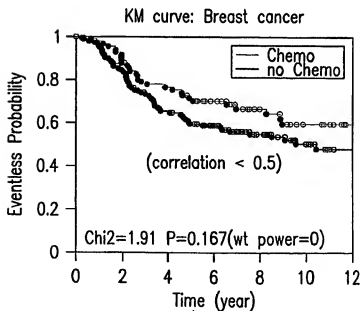


FIG. 10B

13/15

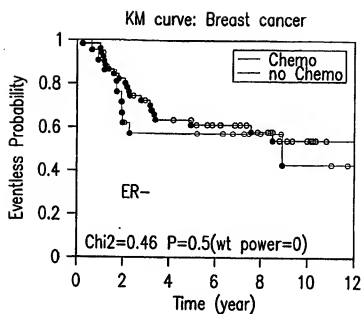


FIG.11A

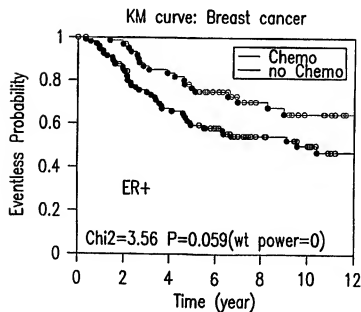


FIG.11B

14/15

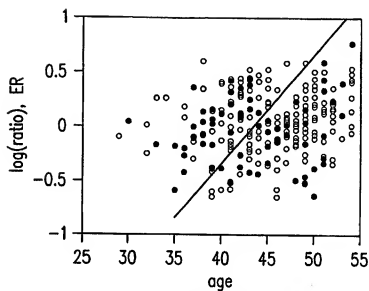


FIG.12

15/15

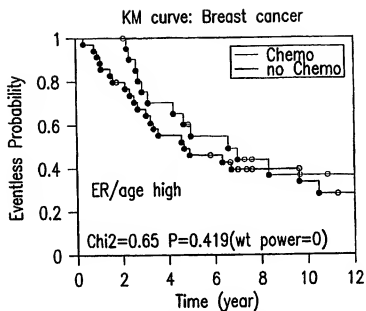


FIG.13A

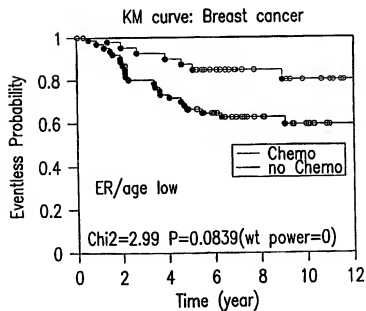


FIG.13B

SEQUENCE LISTING

<110> Rosetta Inpharmatics, LLC

<120> Methods of Predicting Chemotherapy Responsiveness In
Breast Cancer Patients

<130> 9301-249-228

<140>

<141>

<150> 60/650,365

<151> 2005-02-04

<160> 425

<210> 1

<211> 4946

<212> DNA

<213> Homo sapiens

<300>

<308> AB032969

<400> 1

cagcctcagc	ccccagatga	agatggggat	cacagtgaca	aagaagatga	acagcctcaa	60
atgggtgggtt	taaaaaaggg	agaccctgtca	gttgaagaag	tcatgaaaat	taaagcagaa	120
ataaaagcgtg	ccaaagcaga	tgaagaacca	actccagccg	atggaagaat	catatatcga	180
aaaccagctca	agcatccctc	agatgaaaaa	tattcagggt	taacacgcaag	ctcaaaaaag	240
aagaagccaa	atgaagatga	agtaaatcag	gactcggtca	aaaagaactc	acaaaaacaa	300
attaaaaata	gtagcctcct	ttcttttgac	aacgaagatg	aaaatgagta	agtgtaaata	360
ttttgaattt	agtctacttt	gaaagtatat	ggagtgttca	ttaaaatcac	attttttcct	420
attataaaga	tactacaagt	tttttataga	aagtttagga	aatagagaaa	aaaatttaat	480
aaactacatc	tattcatcaa	taccctctcg	acttaaaatg	ccaactctat	agaaattatgc	540
tagtattaac	attttgttat	ttcccttggt	tggttgtata	tatatgtaaa	ttatatatttt	600
aagcaaaaaa	cattttttgt	gtgtaaacaa	aattttataa	atacaactgt	attgcaaatg	660
ttctttgtcc	tgctttccac	ttgacattgc	attatgagta	ttcttccagg	tcagttaatt	720
tcaaaaacct	gacattaata	gctacagata	atttcataaa	catctcatgg	tatctttttc	780
attagcaata	gctccacttt	gggtggggga	gatgataatg	tgcccttgta	aaaataacct	840
cccaactcct	gctaagggtg	gccatgagac	tcagctctgg	caagttaaga	aatacagggtg	900
gaattctcgt	tgataaaagt	gctgggtttt	ttgttacaaa	aggacagact	tggtcaaacat	960
gagcctttgc	ttcttatcttt	tcactctact	tggaagtcag	agataaaaac	tgagtccag	1020
agccactttt	aggcataaag	aaggcagcca	tgtgctttgg	gtcatgttag	taaaaagact	1080
cagagcttgg	ctccttgctg	acatgcctgg	aggagctgct	acaccagctt	ggattgctga	1140
ctctcgactt	cttggtagtg	agaagaataa	acactgtgct	taattagccc	ttggtcaggt	1200
ttcttttata	tcgagccaaa	tgcaactcta	agtaatacaa	taaaactgtg	gtcaaaactgt	1260
tactgtggga	gggtgtccag	gttcttggca	ttttggacaa	ataattgaac	aaaacgcaca	1320
aagcaatgaa	tatcctctag	aggtttgcca	ttgtgttaact	ggcgtaaccc	ctgtgttaaat	1380
gaagtgtggg	cccgtagcct	gtctgattgg	tcgacagaagt	gaccaatcag	aggctggaat	1440
gaagtacaaa	agttatactc	ctgtgtaaat	gaggacttgg	ctatgacca	gtctgattgg	1500
gtgcaggagg	ggaccaatca	gaggcacttt	catttttcat	ctgcaatgca	gaaaaggcaa	1560
ggggattgca	aaggggatag	cctctgatcc	ttttgttact	taggtatgga	gaggtggggg	1620
tttcttttgg	attcagttct	aggaagtc	tgtaaatcag	ccttaggttc	cctgtctcca	1680
gaccctatcc	tctctgctca	ttttccccc	gagagacgtg	atcctcgtaa	atctttatgg	1740
gaggtcgaga	gactgagggt	ctttctctgt	taactgtctc	atgctaactt	gggacacagt	1800
ccctaccctat	ttggagatcac	gtaaactctca	ccctgcttgg	tctaggggag	acagggtgat	1860
ttcttgatgg	ccggtgtgtg	cttctcctga	aactgggtcg	aaatcttgtc	atcatgatcat	1920
ctaacttggg	ggctctctag	caaaaggaaa	tggatttggg	taaaagattt	aacagatatg	1980

gtccaaaac caaggcaaat ataatacatta ataattgggtt ggccaaggga gggagccatg 2040
aaaccaaac tagtgccctt taggtgcccc agctgtgtgc atattttaga ggccacgtca 2100
gctagttttc aggtgtgtgc ccttactaat cctgatttgt tgacatcaaa acagacttct 2160
tctcttagga aaatacataa gccacctgtt tcagcagtta ggagatctgt tccctcttga 2220
tttgcgaag cgaccactgc caaggagcct atccgaattt gtaaggtagc aatactttga 2280
gcaatgttat gcaagtcttc cctagctctc ttggacaagc gttgttaata gtagaggaa 2340
gttgoaatct cgctaaactcc cattcctacc tctgctgtta ttccatgccc ttgtgtctgg 2400
tgggtgcagt taaaggatata atgagggatt ggttgttggg gattatatta attaggga 2460
atcaaatatt tctgtctcca gtctaccact tcaccacaag acaaatcaca gcagaaacca 2520
ctcaacttca aaataaacctg cagtcccata tactgggccc gattaccaca acaagtgcga 2580
acaagaatca ttgtccatat agactctcct agattggcct tgctagaaca ttccacaagg 2640
ccatttcagt caaagtcctg agaaagtaac cggtttcaat tgtgccctat tacaaaagaa 2700
aacgtgtgtt ttaactttat acagacaaat gccatgaatt aagaattatc ataaatagt 2760
tacaatttct ggagaaattt gaatactcaa tacacttaaa gtgtatttca aggctataaa 2820
tagctcaaaa taaaagattt attcagactc tgaaaaaaca gttcatttca gcaattttc 2880
aaacaacaaa agccatacaa attattctcg tcttccatta gtccatttca tccattgtaa 2940
tcaactcctg ctctacttca tattcatctt tatgaacaca tcagccttcc aattagtggc 3000
ttgggaagtt tctgtctaat ccaatggcac actctccaaa gttaccagaa acctgcattc 3060
aagagctctt ttcatgaact ccaagatgga aagccttgga ctgtagctga ttataagtc 3120
ctttttttt ttgagaaagga tcaagcaaaa acatcaatta tggatgacaa aagtccttaag 3180
acagccataa agacacagtt gacaatgtgt gctatttctg tgccctacaa caatttaaca 3240
taatcattac aacatattat aagacatatc tctatcagta taaccacaag gaagctaaac accacctcac 3300
acacattata acaacaatac taattcaaac attacaaata agcctaata aagcctaata 3360
acttgacaat gtttctgtga taattcaaac attacaaata agcctaata aagcctaata 3420
tgtcactctt gaactcagg aagcctaata tccaaaaagt tagtttaagg tcaaaagttt 3480
ttgaattaac ttttttccat tagtatggct atactttct tactaatttg taagttaagt 3540
aattttacaa tttttttttg ttgttctgtt tcccaacct tatgtcagat aagaattcac 3600
ccaggccaga cacagtggct catgcttgtt gtcccagcac tttgggaagc caaggtggga 3660
gaattgcttg aagccaggaa tctgagccca gcctgggcca caaagcaata cccctatctc 3720
tcaaaaaaat aaaaaatagc caggtgtggc gacacacacc tgggtgccca gctgctcggy 3780
aggctgagcg ggaggaatgg ttgggcccag ggttccaagc ctgagctgag ctgtgattgc 3840
gccactgcac ccagcctgg gcaacagagt aagaactgtc tttctcttgt ttttgtttat tcacattaac 3900
aaataaaatt taaaaaaga attaccata actttttctt caaataatca attgtctcat 4020
aaccatgtgt ttctcattta ttgaaaggc catctagtga gagatttctc caaatgttgg 4080
ggtagggaa gggaggggaag cactttaaag tctgagcctt tagaggtgat tctcaagac 4140
cctgcttaat cctaacaatt ttctcatta gtaaaagta gcccaaatg ggggtctgtt 4200
aagatcctta ccagccacat ccacttgaaa ttatgaattt caaagtatct tacaatttg 4260
gtgccacatt atctttttta agttttttt gttttgttt tttgagacag agtctgcctc 4320
tgtcaccocgg gctgagtgcc agtggcgcca tctcagctca ctgcaagctc cgcctctcgg 4380
gttccaccaca ttctctggcc tcggcctccc aagtatctgg gaactcagtc gcccccaacc 4440
acgcccgggtt aatttttttg tattttaaat agagacgggg ttaccatttg tttagccagg 4500
tgggtctcaat ctctgcacct catgatccac ctgctctggc ctcccaagt gctggggata 4560
caggcaggag ccaccgcgcc tggcctttt ttaagtttt aagtacctat aaagaacct 4620
gaaagtgat gtgtgtggat gagctaggaa gactgaaat gctcttctct taaatttaac 4680
aaattatccc tgaagccatt ctgcaaatct gtctttaat tatactcat tggattagaa 4740
gcaggggttt ttccocctaa ttgtatcat tctatatgt ttatatgtac caaactacac 4800
tgttttaatt gctgtaaat ttaatatgtc tttagtatcg ggtgtgggaa tctgtgaagc 4860
atggagtttg tgttatccac cactgtattc tcaaatata gaagagtatc tggctacta 4920
atgtgcacaa aaacatagtt aaaaatg 4946

<210> 2
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> AB032969

<400> 2
taatcctgaa gccattctgc aatactgtct ttaagtata ctcaatttgt atagagcca 60

<210> 3
 <211> 1007
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF005487

<400> 3
 gaatacagaa tgtgggcaaa ctgcttctg tgcggcgccg cagaaggttt gctgagggca 60
 atcactccct ggtgccgggc tcottgaggt tatgcactgg gacatctaga gccattgtt 120
 tgaggaatgc agtcttgcaa gcctgctctg gatcaagcca cagactgaaa caccgccgaa 180
 gagcaagcac gtttcttgga gcaggctaag tgtgagtgtc atatctcaa tgggatgaag 240
 cgggtgcagt acctgaacag atacatccat aaacggggagg agaacctgog ctctgacagc 300
 aacgtggagg agttccaggc agttacggaa ctggggcgcc ctgtcgcgaga gaactggaaac 360
 agccagaagg gcatcccgga ggagaagcgg gacaagatgg acgactactg cagatacaat 420
 tacgggggtt ttctgagagc tcacagtgca gccgcgagtc catcctaaag tgactgtgta 480
 tctcgcaaaag acccagcccc tgcctcacgc caacccctcg gtccgctctg tgagtggttt 540
 ctatccaggc agcattaaag tcaggtgggt ccagaatggt caggaaagaga aggctcggtt 600
 ggtctccata ggcctgatcc agaattggaga ttggaccttc cagaccctgg tgatgctgga 660
 aacagttctt cggagtggag aggtttacac ctgccaaagt gagcatccaa gcgtgacgag 720
 cccctccaca gtggaatgga gtacacggac tgaatctgca cagagcaaga tgctgagtgg 780
 agtgcggggc ttgtgctgg gcctgctctt ccttgggaca gggctgttca tctacttcag 840
 gaatcagaaa ggacactctg gacttcagcc aacaggactc ctgcgctgga ctctgagct 900
 gaagtgcaca tgaccacatt caaggagaa cctcttgcca cagctttgca ggatgaaaag 960
 ctttccactt tggctcttat tcttccacaa gagctctctc aggacca 1007

<210> 4
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF005487

<400> 4
 ttctcaggat gaaaagcttt cccacttggc tcttattctt ccacaagagc tctctcagga 60

<210> 5
 <211> 3200
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF026941

<400> 5
 caggaaaggc catgaagatt aataaagatt tggactcagg gcaaatattt acttagtagc 60
 aataactcaa agaattactg ttgaataaat aagccaatta agcagccaat cactgtactat 120
 gcggatgcac acaaatgaaa cctcacttcc aacctgaaga cattcgcaaca tgagttaact 180
 agagggacct gcaggaaagc gtgagagaaa cataaggctt atgogtttaa tttccacacc 240
 aatttcagga tctttgtcac tgacagcagc actaagactt gttactttta tatagttaag 300
 aagaacaaag ctgagcgcga tgactcagcg ctgtaagcct agaactttgg gaggcctaag 360
 cagcgagact gcttgagccc aggagttcca gaccagcctg ggcaacatgg caacacocca 420
 tctctacaaa aaaaatacaag aatcagctgg gcgtggtgat gtgttctgtt aatctcagct 480
 actcggggagg cagaggcagg aggattgctt gaacccggga ggcagaggtt gtagttaagg 540
 gagatctcgc cactgcactc cagtctggac gacagagtga gactcagctc caaataaata 600
 aataaataca taaataataa gaaaaaataa aagctgtctt ctctctttcc tctcttttgg 660
 tctcactctg ctctgtccca ggcactctgc cactctgagc acagtgtggg tctgtctttt 720
 gctgggaagt tcttgagtgt gttcaggcaa cctctgagct ctctgtggag gagcctgggc 780


```

ccgcgtgttct gctggctgag ggcaaccttc tgctgtctag ctaccaagag gagaaagcag 840
cagctggtgcc tgagagggcc agatgagacc aaagaggagg aagaggacc tcctctgccc 900
accaccocaa ccagctccaa ctagctacttc actgccagtg gcaactacaa atgagggttc 960
tgtttccaca cagccaaaac atcctttgtg ctgccccttg aggaagcaaa gagaggattg 1020
cttttgtcta aggaagctgg tatggagaag atcaactttt cagggtggaga gccatttctt 1080
caagacgggg gagaatacct ggccaagtgt gtgaggttct gcaaaagtaga gttgcgctgt 1140
ccacgcgtga gcatcgtgag caatggaagc ctgatccggg agaggtggtt ccagaaattat 1200
ggtagatttt tggacattct cgtatctccc ttgacagagt ttgacagaga agtcaatgtc 1260
cttattggcc gtggcccaag aaagaagaac catgtggaaa acctcaaaa gctgaggagg 1320
tggtagaggg attatagaat ccctttcaag ataaattctg tcattaatcg ttccaacgtg 1380
gaagaggaca tgacggaaca gatcaaaagc ctaaacctgt tccgctggaa agtgttccag 1440
tgctctttaa ttgaaggtga gaattgtgga gaagatgctc taagagaagc aqaagattt 1500
gttatttggt atgaagaatt tgaagaattc ttggagcgcc acaagaagt gtctgtgtt 1560
gtgcctgaat ctaaccagaa gatgaagac tctacotta ttctggatga atatatgcgc 1620
ttctgaact gtgaaaggg acggaaggac ccttccaagt coactctgga tgttgggtga 1680
gaagaagcta taaaattcag tggatttgat gaaaagtgt ttctgaagcg aggaggaaaa 1740
tacatatgga gtaaggctga tctgaagctg gatgtgtaga ggggaaagt gaacagagact 1800
tcaacacacc agtgggaaaa ctccctagagt aactgccatt gtctgcaata ctatcccgtt 1860
ggattttccc agtgggtgaa aacctgattt tctgctgcac gtggcatctg attacctgtg 1920
gcactgtgaa acacgaataa ctgtgtagc aaatcctgag acaatggaaa accatttaact 1980
ttacttcaat ggcttataac ctgtgtgtta ttgaaacagc acttctgttt ttgagtttgt 2040
tttagctaaa aagaaggaat acacacagga ataatgacct caaaaaatgt tagataaggc 2100
ccctatacac aggcactgac atttagctca atgctcgctt ttgaagaaat agctcctagt 2160
gatattctgt ggggcaatat ttaatttga tttgattttt taaaaaatgt ttactgcga 2220
ttctatattt tccattttga aactatttct gtgtccaggt ttgttcattt gacagagtca 2280
gtattttttt ccaatttccc agataaccag ttttccacat tgagacataa caaatgatct 2340
gcctcaatta ttctgtctgg ttataatgct tttttttttt ttgtcttta tgcattgca 2400
gtctgttact ttttactgtg atgtacagaa atagtcaaca gatgtttcca agaacaatat 2460
atatgataat cctaccaatt ttcaagaagt ctctagaag agatacacca tggaaagacg 2520
cgctgtgtga gccacgcccc cgggtgctgt tccatgaatg cctgagcaa acaaaagata cgctttccat 2580
taactgttgt gtccctttct ctccaaagat cctgactacct ttgttcgct atcatctggc 2640
ttgatgtagg agttgacatg gaggcagtgct ttgcattgtt ttgttcgct atcatctggc 2700
cacatgaggg tgctcaagcaa aagaatagga gtgtagtga gtgactggtt ggccctacat 2760
ttctgagaag tgacgttaca ctgggttggc ataatgatac ctaaaatcac gctggaaacct 2820
tgggcaagga agaattgtgag caagagttaga gagagtgctt ggatttcatg tcagtgaagc 2880
catgtcacca tatcatattt ttgaatgaac tctgagtcag ttgaaatagg gtaacctcta 2940
ggtcagttta agaagagtca gctcagagaa agcaagcata agggaaaatg tcaogtaaac 3000
tagatcaggg aacaaaaatcc tctccttgtg gaaatatccc atgcagtttg ttgtataaac 3060
ttagtatctt attgcctaaa aaaaaatttc ttaattctgt ttcaaaaaag caaaatcatg 3120
gaaaattttt gttgtccagg caaataaaag gtcattttaa tttaaaaaaa aaaaaaaa 3180
aaaaaaaaa aaaaaggcca 3200

```

<210> 6

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF026941

<400> 6

atttttgaat gaactctgag tcaagttgaa tagggtacca tctaggctcag ttaagaaga 60

<210> 7

<211> 1799

<212> DNA

<213> Homo sapiens

<300>

<308> AF035284

```

<400> 7
gcttgaaccg gggaggtgga ggttgacgtg agctgagatc acgccattgt actccagcct 60
gggcgacaga gcaagactcc atttcaaaaa aaaaaaaatc caactcatata 120
aaaggtgagc tcagctcact ggtccatttc tcagtggtct ctccatcctc atttgcgaac 180
ctcagagggg taaggcagtt gaacctgatg agcaagaatt ataacagcaa ggaaacatta 240
atgcttagaa ttctgagatc cagcacaaact cagctctgtg gagctcagct cgctgccag 300
ggataggtat gacctatgct tgcccttaggc tgctgggaga tgccattctc cagtttcaga 360
agcagggcag gcaaaaggtca agactgtggt attggggctt ttgggctctg aaggatcctg 420
gaacctcgtg ttttggttta ttccctccag ggtctaaga gaacaagagg tgctagctct 480
taccaaaaa gctggttagag agagttgctg gctattttaa agctctttc atcttttaat 540
tcacctcttc ttttcacctc ttaaacactc cctcaggaa acgaactctc taggaactgg 600
ggctcttttag ctccataaagc aagtgcagc atgggacaag ttagtctttt ctccatagaa 660
acaaagggga tgcccagtggt tttccctttg ctcccaacc taaaatttca agtttaataa 720
aatagcaatt agcagaagtg accaaaattg gagataatta tcagtcatga ggaaagacac 780
agatttcggt catataaagt gtaagggcta taagtatgaa cttctataa cctaaatgat 840
gttatagaat tatttttgag caggagcaga aagatttaaa atgatcact catacttcta 900
aatcagaat aggaagatta aaaccacaga acagtttgg atttctattg ctggttagcta 960
ggtatcttac tctgtccact ctgttccaag tatctaaact ttctggaaac caaataggct 1020
ttagaagaga ttatctcata ttctatcag tataatacta aaatgtaact tttaaatcat 1080
ctggttttta aaagataaac agtttagccc atctctccag agagcaaaa taggaatatg 1140
actcaggagc ctccatgggc ttatcatcag cctcacacc cgcttcccc tccaaaccac 1200
agcctttgct tcagggtggc aggtattact atttgccctc tcagcagcat ctactctagg 1260
catatgtatc attttagaca ctgggagaag aagacctcaa actaggaga aaagacagag 1320
ctccactcta gttttgggag gggatggcag acagtcaagg agatgagcgt cctaaggcat 1380
ttggggatag ggtccagatgc acccccatg gagaggtttg tcaacacaaa gacatggaag 1440
gttagaggtt ttgtaacaaa aagacatgga aggttaggtt tgtaacaca aagacatgga 1500
agattagagg ttgtcaaca caaagataga ggaagaatgg gctgcagaag atttagatgt 1560
tttccatttg ggcacatttt acttagctgg agaactaggt ttaaaacag ctgggtagga 1620
aaattagaag caagctggat gcagtggtc atgcctgtaa tcccaacact ttggggaggt 1680
ccaggcagga ggaactcttg gcccaggag gtcaagctgt cagcagcgtg agatcacacc 1740
actgcactcc agcctggggg gatagaacaa gacctgtct caaaaaaaa aaaaaaaa 1799

```

```

<210> 8
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> AF035284

```

```

<400> 8
caaaaaagaca tggaaaggtta ggtttgtcaa caaaaagaca tggaaagatta gaggtttgtc 60

```

```

<210> 9
<211> 1380
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> AF052162

```

```

<400> 9
gtcaaaaggat atttatttat aggccttttt ttttttaata tagaatctga ggctgttttg 60
gcttttgactt aaatttccat caggcctctc tccagcaggt aatccctctc ctcccgctgg 120
gtccctgggg gaggtgtgaa ctcaagggcc tagcccoaaa acactttttc gctttttctt 180
aatccttttc cagtcctctc tttttttata aactgtggca ttgtaggtt tctgtttcgg 240
cataacgtaa tccatttcac tgtagctcaa agactcagc gaggttgat attgttcaaa 300
tgagcagggc ccgagctgga agcgcaaggc agcccgccgc gtcccgctcc cctctggccc 360
tcaggccagg tccctgctgg aagcggtcgc atctctctct cagccctggg ttccatgggt 420
actggcgtca ccgagccacc cgagtatggc gaccttctct gcagagagag tgaccgaggt 480
cttttgcttg tggaaaggaga cgctgggctg tgcggtgcgg aggggtgatga ggatgtctgg 540

```

tgacagccgt	goggacacca	ctoctctctg	cagcaactgcc	tccacagccc	agggctcgccg	600
gcacatccca	ctgagagcgg	gggtcctgcc	ccatcttaga	gtcaaaggca	gaggggcttc	660
caggccctgg	atgggggtatt	ttggtgtcac	ctgaagtcoc	tctgacatca	ccttgtttca	720
tcatttttta	tgacagaatt	agaaacccat	ccttcaagca	caataatcat	cacagacttg	780
agtttgcctc	ctaaagcaaaa	ggctccgggt	ttgtttggaa	aatttttttg	atttctgaaa	840
gaatttgatt	tttatattttg	gggcattctct	atagaaagtg	accaccaagg	ccagtaagta	900
cggggaaaaa	tgtttacttaa	cttctcaga	gattcgtgat	acgcgtttct	ccactgacag	960
acatttaaaa	acaaaccttc	gctcgttttc	aatcaatcac	ctcgacttgt	tttttagcat	1020
ggcactgccc	agcaggacag	acaggggatg	agtaaaccca	agtaaatctc	agggctcttg	1080
cgtgttttga	acacagaaga	atcctagtgc	agccttttgt	agctaacagt	cactgatatt	1140
ataattggag	aatgcgtaaa	gattcatttt	tcaaggagaa	gagcctgcaa	atggccaatg	1200
aaggaggtaa	ataaactaag	atattccgag	ggaaggagcc	caggccacct	cctctccgca	1260
ggtctgcaga	tgaagggttt	tttgaatgaa	atgccactgt	gcattttcag	aaaaaaaaat	1320
ctctgataaa	cagactttga	atggaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1380

<210> 10

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF052162

<400> 10

cagtaagtac gggaaaaaat gtttactaac ttcttcagag attcgtgata cgcgtttctc 60

<210> 11

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> AF055033

<400> 11

ggggaaaaa	gctagaaa	agctgcaaa	cagtggtggc	tttttccct	tttttgcct	60
ttttcttacc	ccctctccg	ttttcacct	tctccgact	tgcgtagaa	cctgcgaatt	120
tcgaagagga	ggtggcaaa	tgggagaaaa	gaggtgttag	ggtttgggtt	ttttttgttt	180
ttgtttttgt	tttttaattt	cttgatttca	acattttctc	ccacctcttc	ggctgcagcc	240
aacgcctctc	acctgttctg	cggcgcccg	caccgctggc	agctgagggg	tagaaaagcg	300
ggtgtatttt	agattttta	caaaaatttt	aaagataaat	ccatttttct	ctccccccc	360
caacgcctac	tccactgc	ccgatctcat	tatttcgggt	gttgcttggg	ggtgaacaat	420
tttgtgcttt	tttttccct	ataattctga	cccgctcagg	cttgaggggt	tctccggcct	480
ccgtctcact	cgtgaacctg	gcgtgcctct	gcttccccca	acctgtttga	aggtcttaat	540
ctctgcaact	gggacctgct	cgcaggccac	ccagccctcc	acctctctct	acattttttg	600
aagtgtctgg	gggagggcac	ctgctctacc	tgcagaaaat	tttaaaacaa	aaacaaaaac	660
aaaaaaatct	ccggggggccc	tcttggcccc	tttatccctg	cactctcgct	ctctctcccc	720
accccgaggt	aaaggggcg	actaagagaa	gatggtgttg	ctcacccggg	tctctcgct	780
gctggcgccc	tatggcgggc	cggccacagag	cctgggctcc	ttcgtgcact	gcgagccctg	840
cgacgagaaa	gccctctcca	tgtgcccccc	ggctccctcg	gtgtgcagga	tggtcaagga	900
cccgggctgc	ggctgtctga	tgcactgcgc	cctggccgag	gggcagtcgt	ggcgctcta	960
cacgcagcgc	tgcgccacag	ggctgcgctg	cctccccgcg	caggacgagg	agaaagccgt	1020
ccacgcctct	ctgcacggcg	cgggggtttg	cctcaacgaa	aagagctacc	cgagcaagt	1080
caagatcgag	agagactccc	gtgagcacga	ggagcccacc	acctctgaga	tggccaggga	1140
gacctactcc	cccaagatct	tccggcccaa	acacaccggc	atctccgagc	tgaagctgga	1200
agcgtgaaag	aaggaccgca	gaagaagct	gaccagctgc	aagtttgtcg	ggggagccga	1260
gaacactgcc	cacccccgga	tcatctctgc	acctgagatg	agacaggagt	ctgagcaggg	1320
ccctcggcgc	agacacactg	aggtctccct	cgaggagctc	aaagccagcc	cacgcctggt	1380
gccccgtgct	gtgtacctgc	ccaattgtga	cgcagaaagg	ttctacaaga	gaacacagtg	1440
caaaccttcc	cgtggccgca	agcgtggcat	ctgctggtgc	gtggacaagt	acgggatgaa	1500

gtgcgcaggc atggagtagc ttgacgggga ctttcagtgc cacaccttcg acagcagcaa 1560

cgttgagtga tgcgtccccc cccaaccttt ccttcacccc ctcccacccc cagccccgac 1620
 tccagccagg gcctccctcc accccaggac gccactcatt tcactctcatt taagggaataa 1680
 atatatactc atctatttga ggaaaaaaaa aaaaaaaaaa aa 1722

<210> 12

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF055033

<400> 12

tccacccagg cagcgcactc atttcacttc atttaaggga aaaatatata totatctatt 60

<210> 13

<211> 1411

<212> DNA

<213> Homo sapiens

<300>

<308> AK001166

<400> 13

aaacaaagag atgccacccc tgtgtgatgg ctttggtagc cgaacactga tggttcagac 60
 attttcccggt tgcattctgt gtccaagga tgaagtggac ttggatgagt tattagctgc 120
 tagattggta acgtttctga tggacaatta ccaggaaatt ctgaaagtcc ctttggccct 180
 gcagacctct atagaggagg gtgtggctca tctaagaaga gtccagataa aataccaggg 240
 agctgcatac gatatacatt tatctgctcc atcatttttg cgtcaaatga gtccagagga 300
 atttgaatat caaagatcat atggctctca ggaacctctg gcagccttgt tggaggaagt 360
 cataacagat gccaaactct ccaacaaaga gaaaagaag aaactgaagc agtttcagaa 420
 atcctatcct gaagctctac aagaacgatt tctcacacca gaaagtgcag cactctctgt 480
 tcctgaaaaa cccaacccga aaccacagct gctaagtgg gcaactaaga agcctttcca 540
 accatttcaa agaactagaa gttttcgaat gtaataatac ttccacagca acaggtgcta 600
 gagaccactg ttgtgttttt gagtgaatgg tggtaggag aaagactttg gtggtggag 660
 aagaaaaagc ataaaaaaa gactactgaa atatagataa agattgcctt agtttttaaa 720
 aatgttttgc cattagtatt ttataaaac tcaatgctag ttttaagtgt ataaatttgt 780
 taaaatttat gagtcaata tatagtata atgttaacat gtttgaatt gctacagaat 840
 taaagggtag ttttactctc gtgcttcttt ttcatgggt tttattaaat aattgtgtat 900
 atacatccta gctactgata tctttattat agccttaaga ctttaatttta agtctaaaa 960
 atagcgtgta tacttgaata agaaagacac tgggtactgt tactgtgatg ctattgactt 1020
 agtagccaat tatcatttct cctgtataaa ttccagtttt tattgtgcga cataaatttt 1080
 ttaatgtctt atatttgtat agctatgtct tttattgcag atttatggg tgttatgaca 1140
 gattttacta aaagctagtgt tttataaaca tatatttag ttgatgttta cctataaagt 1200
 gagttagatt tcactgcctt gcaatgggat aatttcaagt tttagctaaa atggaagt 1260
 gaactggata aattcttttg gtaccccttg accctgtatt ctaagtcaaa tgcgaatggg 1320
 ttaataaaaa tgagactact tctttataaa ataattttt atccttttga aagtaagtga 1380
 aatgtaataa aacttatttt ttttaaaat g 1411

<210> 14

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AK001166

<400> 14

acccttagac ctctgattct aagtcacatg caaatgggtt aaataaaatg agactacttc 60

<210> 15
 <211> 2352
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL049367

<400> 15
 ggcaaacccc ttttaaaatc taatgtctgg gctttgagta ttagctcatt taggggtggac 60
 aaatgcatta ctgtttttcaa actgctcaca ttattcagat atttctccaa gttgctatct 120
 actcagcctt atgaatgcco ctgccttttc taaggccatg tgaataatcac ggcaactgcc 180
 ttatgccttg gtcatctgct ttttcgttct gcgatatgcc cagttcccaa atcaattata 240
 ggtaacctgt taggagagag gaagatttta cctctcaaa ggtgagattt gaaatttaca 300
 ctaaaaagac aactttacat ttaatgcttc acttaagtga acattctttt ttttataagt 360
 ctatttttct actcagtttc agaacactaa tctgatttcc actcgtattt ttaacgtttc 420
 ttttaattat tataatgtag ctcttttcaa aatattttca tgaataatta cttttattat 480
 accattatgt gcatgttatt ggtagcaggg atagtttatt atttagtact gaaacatgct 540
 ctctttaccta acagtaaaac agtatgtttt gatataatc tggtaatatg ctatagtggt 600
 taagaatatg aactgaggtc ccaggagatt tcattttatt caccctggtc agatacaata 660
 aaggctatga gtataaatac ataacttctc aaccagggtg agggcatgtt catgaatata 720
 aaatcctttt atgtctggac caagagagga aaagtgttag ctaaatgttg atttacttat 780
 aactagacgt ctatgtgaga aaatatatgt atacatatat atgatatgca gaagtcactt 840
 tttttatcag gctttattct ccttacaagg ccacagttta actgtctgca acagttgggt 900
 tatgttaagt atagacaaat acccagtggt tgttactttt tcoactacc actgtaatga 960
 taactcttct cagctatata catgcaactt ctggcttoca ttccatgaa gctgtttcaa 1020
 tatattcagt ataactgttc cttaagtctg ctctgttaa cagtgatctc tttctttttt 1080
 tcaattctat actctcatta gttcatcata aatctgtcca gttgagacct caggaccacg 1140
 gcatgatttc atgactcoga agtattttac agaaacattt ttaataaag ggaatatatt 1200
 tatataccag atggttcaca agtgatggct catagctagt tttttttt ctctctaaaa 1260
 aatgtccagg ttttaaaatc atttaccotta ttaaaatgaa aagtgccata cttaactttt 1320
 aaaggaaaga cctgaactgc tttttctcta ttttagactg ttttgtactt tactaactct 1380
 taaactatca ggaaaaaaac caaaacttta taccaatgat ttagtaattt tgaggcatag 1440
 ggtagcttac gtatggagg atgtgcaaaa tattctcttc aaatgccaac ttctcaattt 1500
 ataactaaaa tagtgttata tgactaatc ctctgaattt tgatgtaaga tctatatagg 1560
 cccccaaat gatcgtatga catgccagtc atttctcagt gaaataata caataccaga 1620
 gtacattatg ggttttattg ctctctttta tggtagacct gttaatgggg aaaaaatata 1680
 ctaaatcaaa tagaactcta tatctgtatg ttaaaataga gcaactacct gaagtcagtg 1740
 gctggatgca tagccctgga tcaattccca gctctgtctg ctctgtgtga ccttggacaa 1800
 ggogcttcat ctctctgggc ctctatttct ccatttgtaa acaagtggtc tgcagtagat 1860
 gggctgtgag agcccttctc gttccagatg gcttgggtcc aaagacocca cccctctgct 1920
 ggtcctgcca acgtgttggt gctataagct gcttcagata taacatttgtt ttatctataa 1980
 tgtttgttca ttttaatgct cttaaaaggg ctttttgtta tcaagtgtct ttttcttagt 2040
 tttatggact tgattactgt aataatgtct tgtttttagc catgttaata caaacagata 2100
 tctcttgatg gtattagtaa atttgcaatt gatataatc tgatgagatt tgtgtgttat 2160
 gtaaatattc tgggtacagc atctgtccag cactctatta accataatc tgtgatcatt 2220
 atttggaaat atgtcctatg gaaagaataa aagcatgtac ttcacagcta gcattgtcac 2280
 agatttgaaa gaagtttcat taaaagcaco attgctttct ttcaaaaaaa aaaaaaaa 2340
 aaaaaaaa aa 2352

<210> 16
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL049367

<400> 16
 atttggaaat atgtcctatg gaaagaataa aagcatgtac ttcacagcta gcatgttcac 60

<210> 17
 <211> 1130
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL080235

<400> 17
 ggctgcgcga cccgcgcgc cccgcgcgc gccgcgcgc gccgcgcgc cgcacccgcc 60
 ggggcgcga cccgcgcgc agccatcccc gccgcgcgc gccgcgcgc cgtctggctg 120
 cagggcgcgc cgtctgcatt ctgctgccta gacttcagcc tggaggagct gcagggcgcg 180
 ccggctgcgc ggcgaacgc taagccatt gactccagc tggctggctg ctccatgacc 240
 ctggtcatcg tgggtgtggag cgtggccgcc ctcatctggc cgtgcccac catgcgcgc 300
 ttctctccca acgcatgga acagcgcgc accaccgca gccaccgcgc agccaccgcc 360
 gccgcagtc ccgcaggac caccgcagc gccgcgcgc ccgcgcgcgc cgcgcgcgc 420
 gcggccgctc ctccggggg gccgaccaag tgaccgcgc cgtctctccc tgtgtccgtc 480
 ctgtgtccgc gcccttccc ccggggactc ggccggtgt ctctgtgtg 540
 tagtctatct tagttctct tcccagatg gggccgcgc gagaccgcgc cgccttgaa 600
 aagcaagggt tgtgtgcgc ttccagttcc gaagacaga tgtttaagcc ctggactga 660
 ggggtggatc gcagctccga agacggagag gagggaatg gggcccttc cctctattg 720
 catccctct ccgactcct tcccgcacc cactgcctc agattcatg cagaaatga 780
 ccaaatcctg tgtattgtt ttatatatt aataactgt ttaaatgaa gttttagtaa 840
 aaaaaatca aacaaaaag attaaattg tattgtgta gtaagagaag cctttgtat 900
 ctgaacatag ttgtattga aatttgtgt tttttaatt atttaaatg gggggaggg 960
 catgggaagg attaacacc gatatttgt tacgcgtgaa aatgaacttt atgaacttt 1020
 tccaagtga tctatccag gactggcct ggtgggcgt tcttctgta ctatgtgtg 1080
 ttttggctt ttaatacaga catttccct caaaaaaaaa aaaaaaagg 1130

<210> 18
 <211> 60
 <212> DNA

<213> Homo sapiens

<300>
 <308> AL080235

<400> 18
 ctttgaagg caaggtttgt gctgcgcttc cagttccgaa aagcagatgt ttaagccctt 60

<210> 19
 <211> 2498
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL137540

<400> 19
 gctgaacga cagtctgtc cctgtcagag aatgacctg aacgaagag ctcaacattt 60
 tacacactat gcaatctatg atttcattgt caaggcgagc tgctctgca atggccagc 120
 tgatcaatgc atacctgttc atggcttcag acctgtcaag gccccaggaa cattccacat 180
 ggtccatggg aagtgatgtg gtaagcaca caccagcagc agccactgcc agcactgtgc 240
 cccgttatac aatgaccgac catgggagc agctgatggc aaaaacgggg ctcccaacga 300
 gtgcagaacc tgcgaagtga atggcagtc tgatacctg cacttcagc ttaatgtgtg 360
 ggaggcatca ggaatcgta gtggtgtgt ctgtgatgac tgtcagcaca acacagaagg 420
 acagatttgc cagaggtgca agccaggctt ctatcgtgac ctgcggagac cctctcagc 480
 tccagtgct tgcacaacgt gtccctgcca tcagtagga tccagtgcc tccctgccaa 540
 ctccagtgacc ttctgcgacc ccagcaatgg tgactgccct tgcaagcctg ggggtgcagg 600

```

gcgacgttgt gacaggtgca tgggtgggata ctgggggcttc ggagactatg gctgtcgacc 660
atgtgactgt gcggggagct gtgacctcat caccggagac tgcatacgca gccacacaga 720
catagactgg tatcatgaag ttcttgactt ccgtcccgty cacaataaga gcgaaccagc 780
ctgggagtyg gaggatgcgc aggggttttc tgcaattcta cactcaggtg aatgcgaatg 840
taagggaacg acattaggaa atgcgaagcg attctgtgga atgaatatatt catatgtgtc 900
aaaaataaag attttatcag ctcatgataa aggtactcat gttgaggtca atgtgaagat 960
aaaaaaggtc ttaaaatcta ccaaactgaa gattttccga ggaaagcgaa cattataatcc 1020
agaatcatgg accggacagag gatgcacttg tccaatcttc aatcctgggt tggaaatcc 1080
gtagacagga catgaggagata taagaacagg caaactaatt gtgaatatga aaagctttgt 1140
ccagccatgg aaacctcttc ttggaagaaa agtcatggat attttaaaaa gagagtgcga 1200
gtagcattaa gatggatagc acataatggc acttgtctat gtacaaaaca caaactttag 1260
agcaagaaga cctcagacag gaaactggaa ttttttaag tgccaaaaaa tatagaaatg 1320
tttgaatgca tgggttttat ctaacttata tctctggac ccattgttaa atacagtttt 1380
atttcatgaa gagaaatgaa aacctctaca ctgatattct ttttctatgg gactgattct 1440
gaaattotta actattaaga atattttaat agcagcatga catttagcag taattcatta 1500
agggcagtac ctctaacaag gacgccttcc agcttcagcg atgttactta cgtttgatgc 1560
tacttaaggt aatgaatgac gtttttaagga atccctaacc ctactatcag aaaaggtggt 1620
gtttaaagag ccttctcttg tgtgttaacg atgaactttg gtcctgtaggt gttaaatgga 1680
acctctccat gtgtatatag tatttctctg tataaagcac tttactacct accacttggt 1740
tgtgtgaact ttggtgactg ctgttgaaag aaggaagagg gtgtgtgaga aagcctactg 1800
aagcagcagc actgcacta catgtggaca aaagtgcaca tataaaagaa gttgtgctat 1860
ttaactctga actactggag aaactagggt aagatgcaac cagaagaggag aatagtgtatg 1920
cgtgaagtct cagctttgag ctggaggcta gattccaaga tgacagccat gatgaaactt 1980
tttaaaaaac taacaccagaa gagactttaa aataagagaa agaaatcata aatgtagaca 2040
tatgcttggc taaaggggaa atggacttta aattttaaag agctcatttg caatgcactt 2100
gtatacactt caaaaattat tgtagacaca gaatttgtta tatttttgtg cttagtattt 2160
aaaactgaac attgaaacag ttttctctct gtctcttctt aacagtaata gtcattatat 2220
ttacctgttt ttaacaacaa tgtatgtgat agtcaaaaaa tcacagtttt tcatatttat 2280
tcactctctg taaccacgca taaccactat acatagtttc ttttgtactt gaataacaa 2340
aacatgaaca oagtgccata tgaataattt atatatattg ctttgttaat ttgtttttat atttcatata 2400
ctgtggactt gccaatatat tgaaaaaaa aaaaaaaa 2498
tgtaataaag gaatatgac
```

<210> 20
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> AL137540

<400> 20
tggaggtcatg attccaagat gacagccatg atgaactttt ttaaaaaact aaaccagaag 60

<210> 21
<211> 914
<212> DNA
<213> Homo sapiens

<300>
<308> AL160131

<400> 21
cgcaccgcag gagcaacggt tggctcctgcg gctgtgatgt cgggtgttag gccctggac 60
aagctgcccg gcttgaacac ggccaccatc ttgtctggtg gcaaggagga tgcctctctg 120
cagcagctgg cggactcgat gctcaaaagag gactgcgcct ccgagctgaa ggtccacttg 180
gcaaaagtccc ttcctttgcc ctccagtggt aatcgggccc gaattgacct gactgtgttt 240
gtggttaact ttcacagcaa atacagtcctc cagaacacag aggaagccct gcgccatgtg 300
gatgccagct tcttcttggg gaaggtgtgt ttctctgcga caggtgctg gcgggagagc 360
cactgcagca ttacccggca caccgtggtg aagctggccc acactatca aagcccctg 420

```

ctctactgtg acctggagggt ggaaggcttt agggccacca tggcgagcg cctgtgtgcg 480
gtgtgtcgaga tctgtgtctgg ccactgtccc ggtgtctcag ctctgaacct gctgtccctg 540
ctgagaagct ctgaggggccc ctccctggag gacctgtgag ggtggctggc ccctgggctg 600
ccctctctca tggctctctg ctgactccat aaacattctc tgttgaggat gtccagtcag 660
ggcttgacag gccacaggctc agcccgcctg ggtctgggaag gttccctgca gtgcagctg 720
tgcgcagggt agagctgggc agaagcagcg agggggccca cgtggcgaga ctgtagcccc 780
ctccactctc cacactcact cttgcagagc ctgtgtcttt aagcagctgg cgtgttacat 840
ctccatttaa ggtttccctt gaacaaaagg tctgtggcta aaaaaagttt aaaaatcact 900
gtctctatcc acca 914

```

<210> 22
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL160131

<400> 22
 agctggcgtg ttacatctcc atttaaggtt tcctttgaac aaaaggtctg tggctaaaaa 60

<210> 23
 <211> 4753
 <212> DNA
 <213> Homo sapiens

<300>
 <308> D13642

<400> 23

```

cttcaatcaa gtgaccttcc cactgcagta cacaccagc aaatttgtca tccaccctga 60
gagtaacaac ctattattca ttgaaacgga ccacaatgcc tacactgagg ccacgaaagc 120
tcagagaagag cagcagatgg cagaggaagt ggtggaagca gcaggggagg atgagcggga 180
gctggccgca gagatggcag cagcattctc caatgaaaac ctccctgaat ccactcttgg 240
agctcccaag gctggcaatg ggcagtgagg ctctgtgatc cgagtgatga atccattcca 300
agggaaacaca ctggaccttg tccagctgga acagaatgag gcagctttta gttgtgctgt 360
gtgcagggtt tccaacactg gtgaagactg gtatgtctgt gtgggtgtgg ccaaggacct 420
gatactaaac cccgcactgt tggcaggggg ctctgtctat acttacaagc ttgtgaacaa 480
tggggaaaaa ctggagtttt tgcacaagac tctgtgtgaa gaggtccctg ctgctatttc 540
ccacttccag gggagggtgt tgatttgtgt ggggaagctg ttgcgtgtct atgacctggg 600
aaagaagaag ttactccgaa aatgtgagaa taagcatatt gccaatata tctctgggat 660
ccagactatt ggacataggg taattgtatc tgatgtccaa gaaagtttca tctgggtctg 720
ctacaagcgt aatgaaaacc agcttatcat ctttgtgat gataccacc ccgcatgggt 780
cactacagcc agccctctgg actatgacac tgttgctggg gcagacaagt ttggccaact 840
atgtgtgggt aggcctccac ctaacaccaa tgaatgaat gatgaggatc ctacaggaaa 900
caaaagccct tgggacccgtg gcttgctcaa tggggccctcc cagaaggcag aggtgatcat 960
gaactcccat gtcggggaga cggctgtctc ctgcagaaag acccgcgtga tccctggagg 1020
ctcagaatca ctgtctcata ccactctgtc tggaggaatt ggcatcctgt tgccattcac 1080
gtcccatgag gacctgactc tcttccagca tggggaatgt gctgcgggt ctgaaactcc 1140
ccctctctgt gggcggggacc acctcagctt tcgctctcac tacttccctg tgaagaatgt 1200
gatgtatgga gacctctgga agcagttcaa tccatggaa ccaacaacaa aaaagaaagt 1260
ctctgaagaa ctggacggaa cccacccgga agtgtccaag aaactcgagg atatccggac 1320
ccgctacggc ctctgagccc tcttctcccg gtggggctgt ccagagactc gtgtgtttgt 1380
ttccccaccc accatcactg ccactctggt tctgccatgt ggcaggaggg tgactggata 1440
attaagactg cattatgaaa gtcaacagct ctttccctc agctctctc ctggaatgac 1500
tggcttcccc tcaaatgggc actgagattt gctacaactc tcccaccctg gtacatgata 1560
catgaccoca ggttccagtc tagaacctga gtccccoatt ccccaagacc atccctgcat 1620
tgatatgtct tgaactctct gtctactttt gcacacacc ttaattttta attggttttc 1680
ttgtaaatac agttttctgt aatgttatc ctgtggagg aaggaggcag gctgtgtgtg 1740
gactgggtag ggtatagtat cactcctgag tccactctc ctagaatcta accagaaata 1800
gaaacctagt ttttaaggtg actggcatcc atgtgtctgt tctctggagt gaggatgtag 1860

```


gtgggaggtt	tgaacccaag	ttagagcagg	aagaactgag	tagaactcct	ccttcocagat	1920
accgacttgg	acttgcggca	ctctgtggct	ccccacccc	aggctctgtg	tgggtttctt	1980
gttttttcc	ggttctttt	gotgtgtcta	tgaacaatga	cctcaataac	catgtgtata	2040
ccccacccc	ttccacttgg	gtattgagga	agggtggctg	attctctctc	ctctcttact	2100
ctgaggatgt	tagtatgggg	attttagcat	gaattccagc	tggggagctt	taacagatgc	2160
ccottttact	gattagagac	ctaaagcgat	ctttggctcc	ataggacat	aggaagggct	2220
agtacagaag	aaactagata	ctgcctgcc	cctgagaact	gtgtatatgt	ggggcctgtc	2280
tgcacaccc	atctcaggtg	ggttccagag	ggcctttag	gtataatgag	agcctgttag	2340
gtggaagagg	ccagcttcca	gaaatgttcc	agccaccccc	tgaaatcttc	tctgttttag	2400
ttgtgtggga	agccctcgct	ttccaggctg	tcttggcgcc	ttgaacctgg	aggaagtgcg	2460
ctactgttct	caatacttca	caaatgtaaa	actttcttcc	gtctgcattg	gctcagccat	2520
ctaaattgag	caaatgatct	ggtagagcct	gggttagaat	caggaatggg	ggaatacaat	2580
ctgaacctct	cagagccca	aacagagggg	ctgtgacact	gtgacactgt	ctcctggaac	2640
taagtatctc	ttgaatcatg	acttgggttt	agatcagtc	agagagaccc	aggttttgcc	2700
aggaatcgaa	tccttaata	acatgtttt	ttctcactta	gctcatgaat	tgtcatagta	2760
gacatgagtt	ctgaattaga	ttttgaaaa	ctaatttcag	ggctcatttt	ttcctgtggc	2820
ccctaaatcca	tctctacaaa	ttgtgtgata	ctgacatgca	gtcatctgag	gaactcagcg	2880
tagtactctg	aggagctcct	cgcctctttt	ctaactcaag	tttgactaaa	atacatacac	2940
tccgtacaga	aggtaggggg	ttatgtaaga	aaggaaaacc	taactatgag	actcagaggt	3000
tgtcacacc	gagctctctc	tggaaagtctg	cccatcagct	tgtctgttct	ctgttaagag	3060
gaaggggctag	gacaaggatt	tgggcttgaa	tggtggaaa	ggaattttca	tagtgtgtgc	3120
tgcaggacct	acaaaagttt	aaaattagat	tggatgtgac	tcaatgacaa	gtcccatctg	3180
tgttaattgt	aaggggacct	gattgactcc	tgtgggttga	ttagcaaccc	aggttaatat	3240
agacctctct	ccagcttttg	caaaacccat	cagaggctgc	tgcagaacct	agacagaggg	3300
atctgccctt	gggtttgtct	ccatcctgtt	ccattgtotaa	gcccctgtga	cttggatcct	3360
aggagtgaaa	agtttttagc	tgcctcagct	ttccctgac	cttactggca	gaggtttctg	3420
agatgtttcc	tttggaaat	ctcttgccaa	gaatgcaatt	cotttggagc	aggggggttc	3480
tagttggaat	gttgcttttc	ttggttagtg	taaatgtatt	gctagtgaga	gaggtgcggg	3540
cgctggaaaa	ggctcgtctc	acagggagag	tgtcgtgtcc	cagaatgtgt	gctgttccca	3600
cgctcgtgcc	ttctctgagc	ttgttagagg	aaagccagaa	aggcatcag	atgggatcag	3660
tctgtgtttc	aaattttttt	taattcctaa	gtctgttttt	attttttaat	tttttaaaaa	3720
aaattttatt	agagacagtc	tctctctctt	ctcctagctg	gagtgagctg	gagtgatcat	3780
agctcaactga	ggcttgaact	cctgggctcg	agcaatccac	ctcagcctcc	agagtagggg	3840
agactacaga	tgtgtgccac	catactcagc	tagtttttaa	actttctgta	agacagggct	3900
tccctgtgtt	gcccaggctg	gcctcgaa	cctgacctca	aaaaatcttc	ctgccttggtc	3960
ctccacggcg	tttgagaggc	tgaggcagga	ggatcccttg	agcccaggag	tttgagacca	4020
gctctgggcaa	catgacaaaa	ccccatctct	ccaaaaatcc	aaaaatggc	caggcttggt	4080
gggtgcacat	tgtagtccca	gtaattaggg	ggctgagaca	ggaggatcac	ttcagcctat	4140
gagtttagag	ctgcagtgag	ctgtgattgc	gccactacac	tccagcctgg	atgacagggc	4200
gaaacctgtc	tcaaaaaaac	caaaaaacac	aaacccgtct	ctcggggttc	tggtagcaca	4260
aacgcacatg	actgagtgct	cagggggtct	gagggtgtgc	cgctgacctg	gggctctggc	4320
ctcgggagat	ctgggggacc	tgtgtccta	tatgtatgc	tttgaaagaa	aggggcatca	4380
ttccaagcca	agaggcccca	gagaggccac	cgtggggtgt	tcagggtctt	gtgaggcccc	4440

agttagatcc	tgtggctgtg	cccccatcac	ctccacccc	tctgccctcc	cactagetgc	4500
ccaacggatg	aatcaacgcc	ttggcagagt	tttcacagag	ggccttgca	agagtgtgtg	4560
tgaacctgtg	ggccactggc	ttggggacgg	gtgaggaggt	agcctggaa	attccagcgt	4620
gggcattatt	gtctctgtgc	aagttcaggg	caaaacacag	aatccagttt	tgtcgatcca	4680
attgagaaaa	catttcatga	acaactactt	gtggcatgca	ttggcatcgt	gaataaacgc	4740
cactattgtc	act 4753					

<210> 24

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> D13642

<400> 24

aaaccaggaa tccagttttg tcatccaat tgagaaaaa tttcatgaac aactacttgt 60

<210> 25
 <211> 2591
 <212> DNA
 <213> Homo sapiens

<300>
 <308> D25328

<400> 25

```

ccggacgtg cggtccctt cggcctctc gccatggagc cggacgactc ccggggcccc 60
aagggtcct tgcggaagt cctggagcac ctctccgggg ccggcaaggc catcgcgctg 120
ctgaccagcg cgggggatgc tcaaggtatg aacgctcgcc tccgtgcgct ggtgcgcatg 180
ggtatctacg tgggggccc aaagtgtact atctacgagg gctaccaggg catggtggac 240
ggaggtccaa acatcgacga gcccgactgg gagagtgtct ccagcatctc gcaagtgggc 300
gggacgatca ttggcagtgc cgggtgccag gcccttcgca cgcgggaagg ccgctgaag 360
gctgcttgca acctgctgca gcgcggcatc accaacctgt gtgtgatcgg cggggacggg 420
agctccacgg gggccaaacct ctccggagag gagtggagtg ggtgctgga ggagctggcc 480
agggaacggcc agatcgataaa ggaggccgtg cagaagtacg ctacctcaa cgtggtgggc 540
atggtgggct ccatcgacaa tgatttctgc ggcaccgaca tgaccatcgg caccgactcc 600
gcctcgaca ggcatacga ggtcgtcgac gccatcatga ccacggccca gagccaccag 660
aggacgtctc ttctggaggt gatgggacga cactgtgggt acctggccct ggtgagtgc 720
ttggcctcgg gtgcgactgg ggtgttctt ccagaatctc caccagagga agctggggag 780
gagcagatcg gtgtcaaaact ctcgaggaac cgtgcccgga aaaaaggct gaattattt 840
attgtggctg aaggagcaat tgatacccaa aataaaccca tcacctctga gaaaatcaaa 900
gagcttctcg tcaocgactg gggctatgac acagtggtga ccatctcgg gacgtgcag 960
agaggaggga ccccttcggc attcgacagg atcttggcca cgcgcatggg agtgaggaca 1020
gtcatcgctt tgctagaggg caccocggac accccagctt cgcgtgtgtc actgaacggg 1080
aacccagccg tgcgcctgcc gctgatggag tgcgtgcaga tgactcagga tbtgcagaag 1140
gcgatggaac agaggagatt tcaagatgcy gttcgaactc gagggaggag ctttgcggcg 1200
aaactgaaca cctacaagcg acctgccatc aagctgcggc atgatcagat cccaagacc 1260
aattgcaacg tagctgtcat caacgtgggg gcaccgcgg ctgggatgaa cgcgcggcta 1320
cgtcagctg tgcgcgtggg cattgccgac ggccacagga tgctgcccat ctatgatggc 1380
tttgacggct tgcocaaagg ccagatcaaa gaaatcggct ggacagatgt cggggcttgg 1440
accggccaag gaggtcccat tcttgggaca aaacgcttct tccgggggaa gtacttggaa 1500
gagatgccca cacagatgcy caccgacagc atcaacgcgc tgctgatcat cgttggattc 1560
gaggctacc tgggactcct ggagctgtca gccgcccggg agaagcacga ggagtctgt 1620
gtccccctg tcatggttcc cgctaactgt tccaacaatg tgccgggttc cgattcagc 1680
atcggggcag acacgcacct gaacactatc accgacacct gcagccgcat caagcagtc 1740
gccagcgga ccaagcgcg cgtgttcatc atcgagacca tggcggtcta ctgtggtac 1800
ctggccaaca tgggggggct cgcgcggcga gctgatcgcc catatattt cgaagagccc 1860
ttgcacatca ggatctgca gtccaaactg gcgacactga cggagaaaaa gaagaccacc 1920
atccagagag gccttgtgtc cagaaatgag agctgcagtg aaaactacac caccgaattc 1980
atttccagc tgtattcaga agagggcaaa ggcgtgtttt actcgaggaa gaacgtgctg 2040
ggtcacatgc agcaggttgg ggcacctct ccaatttgata gaaactttgg aaccaaaatc 2100
totgcagagc ctatggagtg gatcactgca aaactcaagg agggccgggg cagagaaaaa 2160
aaatttacca ccgatgattc catttctgtg ctgggaataa cgttattttt 2220
caacctgtgg cagagctgaa gaagcaaacg gattttgagc acaggattcc caaagaacag 2280
tgggtgctca agtaccggcc cctcatgaaa atctggcga acgtatgac cagctatgac 2340
gtgtcggaact caggccagct ggaacatgtg cagccctgga gtgtctgacc cagtcgccgc 2400
tgaatgtgct tgaagccacc gtggaactgc tggttttcta acacttaagt tatttatca 2460
gcactttatg cacttatatt tgacattaat acctaatcgg cgagtgcaca tctgcccac 2520
cagctccagt cgtgctgtc tgtggagtg gtctatgct ttcagatgtg catatgagca 2580
gaattaatta a 2591

```

<210> 26
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> D25328

<400> 26

tattttatca gcactttatg caagtattat tgacattaat acctaatacg cgagtgccca 60

<210> 27

<211> 2573

<212> DNA

<213> Homo sapiens

<300>

<308> D50402

<400> 27

gaatcgccg atgtgaaccg aatgttgatg taagaggcag ggcactcggc tgcggatggg 60
 taacagggcg tgggtcggca caactacttg caccagtgcc cagagagggg gtgcagggctg 120
 agggagctgc cagagcaccc ctacactcc cagagtaact gaagtccgca ttccaatgac 180
 aggtgacaag ggtccccaaa ggctaagcgg gtccagctat ggttccatct ccagcccgac 240
 cagcccgacc agccccaggc caccgcaagc acccccaga gagacctacc tgagtggaaa 300
 gatccccatc ccagacacaa aaccgggcac ctccagcctg cggaagctat gggccttcac 360
 gggctcctgg ttcctcatga gaattgcttt cctggaccga ggaacatcg agtcagatct 420
 tcaggctggc gccgtggcgg gattcaaat tctctgggtg ctgctctggg ccacgctggt 480
 gggcttgctc tgccagcgac tggtcgacg tctgggcgtg gtgacaggca aggactggg 540
 cgaggtctgc catctctact accctaaggt gccccgaccc gtctctggc tgaccaatga 600
 gctagccatt gtgggctcgg acatgcagg agtcatcgcc acggccattg cattcaatct 660
 gctctcagct ggacgaatcc caactctggg tggcgtctcc atcacatcg tggacacctt 720
 cttctctctc ttctccgata actacgggct gcggaagctg gaagcttttt ttggactcct 780
 tataaccatt atggccttga cctttggcta tgagtatgtg gtggcgctc ctgagcaggg 840
 agcgcctctt cggggcctgt tcttgccctc gtgcccggcg tgccgcccac ccgagctgct 900
 gcaggcggtg ggcattgttg gcgccatcat catgccccc aacatctacc tgcactcggc 960
 cctggctcag tctcgagaga tagaccggcg ccgcgcagcg gacatcagag aagccaacat 1020
 ctactctcgt attgaggcca coactcgccct gtccgtctcc ttatcatca acotctttgt 1080
 catggctgtc ttggggcagg ccttctacca gaaaaccaac caggctcggt tcaacatctg 1140
 tgccaacagc agcctccacg actacgccaa gatcttcccc atgaacacg ccaccgtggc 1200
 cgtggacatt taccaggggg gcgtgatctc gggctgcctg ttccgccccg cggccctcta 1260
 catctgggcc ataggtctcc tggcggtggt gcagagctcc accatgacgg gcacctacg 1320
 gggacagttc gtgatggagg gcttctcag gcgtcggtgg tccagcttgc ccggtctcct 1380
 cctcaccgcg tctcgccca tctcgccac cgtgctcgtg gctgtcttcc gggacgtgag 1440
 ggaactgtcg ggctcaatg atctgtctaa cgtgctcag agcctgtgc tccctgtcgc 1500
 cgtgctgccc atctcaagct tcaaccagat gccaccctc atgcaggagt ttgccaatg 1560
 cctgctgaac aaggtcgtca cctcttccat catggtgcta gtctgcgcca tcaacctcta 1620
 cttcgtggtc caactatctgc ccagcctgcc caccctgcg tacttgcggc ttgcagcctt 1680
 gctggccgca gccatacctg gccctcagcac ctacotggtc tggacctgtt gccctggcca 1740
 cggagccacc ttcttgccgc acagctccca ccaccacttc ctgtatgggc tccctggaag 1800
 ggcagaaaaa ggggagaccct ctggctaggc ccacaccagg gctgggtgg gagtggactt 1860
 tatgacgtga ctggcctgct ggtatggagg ggggcgcgtg caggcagcag gatggatgg 1920
 gacagttcct gagaccaccc aacctggggg ctttagggcg tcatctgtaa aatggagagc ctagcgacg 1980
 catgtgatta cctctcgggt ctcatgtccc tctctggcact tgggacaaaa ccaacaaaa 2040
 tgccatggag gttaaagcact ttaacacagt gtctggcact gggcagctgc aggcgtgacc tgacagccca 2100
 aacaaaaaaa ctttcaaaa ggtatttatt gggcgggaca caggctccaa gactggactt 2160
 agggctgggt ggggtagggt ttggagactt gggcgggaca tatctatcta ttatttatt 2220
 gaaatagtgt ctgatgaatg ttaaatatc ttgagtgcat tggagtgacg tggcgcaatc ttaactcatt 2280
 aggaaaaggc tctccctctg tgcgaaggc ttgagtgcat tggcgcaatc ttaactcatt 2340
 gcaacctcca cctctcgggt tcaagcgat ctcttattc agccccggga gtggcgcgcg 2400
 ccacccagcc cagctaattt gtgtatttt agcagagacg gggtttgcca ctgtggcag 2460
 gctggtctcg aactgctgga tcaagtgat ccgccatct ccgtctccca aagtgtggg 2520
 aattacaggc gtgagccacc aaaaaccggc ctgattaaag ttaataaat acg 2573

<210> 28

<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> D50402

<400> 28
tggagggttaa gcaactttaac acagtgctctg gcaacttgga caaaaaaaca caaacaaca 60

<210> 29
<211> 3672
<212> DNA
<213> Homo sapiens

<300>
<308> L27560

<400> 29
acatgtgcat atttcattcc ccaggcagac attttttaga aatcaatata tgcoccaata 60
tggaaagac ttgtctctcc acggtgacta cagtacatgc tgaagcgtgc cgtttcagcc 120
ctcaattaat tcaatttgta agtagcgac gagcctctgt gggggaggat aggctgaaaa 180
aaaaaagtg gctcgtatatt atctacagga ctccatatag tcatatatag gcatataaat 240
ctatgctttt tctttgtttt tttctttctt cctttcttcc aaagggttgc attaactttt 300
caaatagatt cctatagggg cattgaggag ctctctcatt ctgggaaaaa tgagaaaacc 360
catattctcc taatacaacc cgtaatagca tttttgctgt cctcgaggca gagtttcccg 420
tgagcaataa actcagcttt tttgtggggc acagtaactgg atttgacagt gattccccac 480
tgtgtttcat ctgcacccac cgagccaggc agaggccagc cctcgtggtt gcacacagca 540
cgccgctcag tccatcccat tttagtcctt aaacctcag gaagtcacag tctccggaca 600
ccacaccata ttgagcccaa caggctccacg atggatccac ctatgtccac cccagccttt 660
tctttctatc tgaacagaat gtgcattttt ggaagcctcc ctcaactccc atgtctggcag 720
agcaggaggg agactgaagt aagagatggc agagggagat ggtggcaaaa aggttttagat 780
cgaggagaac atgaagatgg atggttccgg ccagagtcga tgtggggagg aacagagggc 840
tgaaggaga ggggctgac tgttccattc tagctttggc acaagcagc agaaaggggg 900
aaaagccaat agaatttccc ttagcttccc caccatagt attttcatgg atttgagggg 960
aaagagagga aaatggggga atgggttgca aaatagaat gagcttaatc caggcccgag 1020
agccagggaa ggtgagtaac cttaggaggg tgctagactt tagaagccag ataggaagaa 1080
tcagttctaa ctggccatgc tttggaaggg acaagactat gtgctccgct gccacacctc 1140
agcctgcaat gagggaactga ggccacagag tctttccagc tcttctcca tctggccag 1200
tccctgcatc ctccctgggg tggaggatgg aaggaagctt gggacaagca gggaaagcat 1260
gatcaggaga tgcgttcaact cggcagccag attccgaac tccacttctc caatgacttc 1320
ctcaaccaat ggtgtgacct gtgactgttc ttttaagctg aagatataca ggaaggggg 1380
cttggcaact ggccaaggag accccttgt gctgtggaca cagctctctt caactcttgc 1440
tcatggcatc acacaggcga gaccgctcc aacaacgaat ttggggctac gaagagggaat 1500
agcgaaaaag caactctgtt tcaactgatg ggaacccat agctatagaa cttgggggct 1560
atctcctatg ggacagttag ggaagcttgg ctggggagac gagaagtgtc caactctcat 1620
gagacaaagg ggcccgatca aggcagccac aaggccttga cctgcccagt cagcatgccc 1680
catctctctc gacagctgtc ccttaaaccc aactcacgtt tctgtatgtc ttaggccagt 1740
attccaaacc tcttccacgt cactgttctt tccaccatt ctccctttgc atcttgagca 1800
gttatccaac taggatctgc caagtggata ctgggggtgcc actccctgc gaaaagactg 1860
agccaggaaac taacagctcc cccacattc ctccagcctt ggacctaat cttagagggg 1920
gctctctctt cctggaactg tctgtgaact tgagcaggct tctgcccctt cgttggctgc 1980
tttctgcga gccatcaggt gggggattag agcctgtgtt aagtgcgccca gactcttcag 2040
tgttccaaag tctgtgctcg cgaacccaaa cctgtgagtc tcttctgcat gcaggagttt 2100
ctctctggcca gctggtcact cccagagaa gctgggacct catggacaca tggaaactaag 2160
ctctccaaat gggagttctg gctgaqccca ggggtgggag atcctgggaa gggaggcact 2220
ggaggaagac ggcacactct ccccatggc aggggtgtgag ggaggcaggt ttggaatggt 2280
gcagatattg caatctaagc aggggtctggt tctctttgac tccaggctgt ctttggccga 2340
ctgtctgctc acccagagac ctggactcc ggaactatca ttgctccgaa tctaagtgtc 2400
gcccactccc atgtctcaac ccacagaagg tcttccatc cctttagat tctgtgctca 2460
ctccaccagt gaggaaagtg cctctgtctt tcccaagac gaggagagat agggagacc 2520
agccaggact gaccctctct cctccagcct gcctgaccc acctggcaaa gacgggcaca 2580

tggggaggaa	gagactggaa	cctttctttg	acagccaggc	ctagacagac	aggcctgggg	2640
acctgtggcc	atgaggggag	gaaggcaggc	gcacgaggtc	caggggaggcc	cttttctgat	2700
catgcgccct	ctctcccaac	ccatctcccc	accaccaact	ctgtggccctc	catggtaccc	2760
ccacaggcgt	ggcctccctc	agaggggtggg	cctcaaccac	ctcgtccccc	caacgaccgg	2820
ttagtggagc	aggggtgccca	cgcaaccgcc	aagccccctc	caaggtggga	cagtaccocg	2880
gacccatcca	ctcaactcctg	agaggctcog	gcccagaatg	ggaacctcag	agaagagctc	2940
taagggaaga	aaaccccata	cggtcagaga	ggatatgtct	ggcttccaag	agaagggagg	3000
ctcgtttttg	caaagtggag	gagggacgag	ggacaggggg	ttcaccagcc	agcaacctgg	3060
gccttgtact	gtctgtgttt	ttaaaaccac	taaagtgcaa	gaattacatt	gcactgtttc	3120
tcacattttt	attttctctt	aggcttttgt	ttctatttca	aacatacttt	cttgggtttc	3180
taatggagta	tatagttttg	tcatttcaca	gactctggcc	tcctctcctg	aaactccttt	3240
ggatggggaa	agggaaagtg	gggagggtoc	gaggggaagg	ggaccocagc	ttccctgtgc	3300
ccgctcacc	cactccacca	gtccccgttc	gccagccgga	gtctcctctc	taccgccact	3360
gtcacacogt	agccacatg	gatagcacag	ttgtcagaca	agattccttc	agattccgag	3420
ttgtaccgg	ttgttttcgt	tgttgttgtt	gtgttttttc	tttttctttt	tttttttgaa	3480
gacagaata	accacagtac	atattactgt	agttctctat	agttttacat	acattcctac	3540
cataactctg	ttctctcctc	ttttttgttt	tcaactttaa	aaacaaaagt	aaacgatgat	3600
aactcttact	ggtgaaaagg	atggaaaaat	aaatcaacaa	atgcaaccag	tttgtgagaa	3660
aaaaaaaa	aa	3672				

<210> 30
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> L27560

<400> 30
 agcaacctgg gccttgtact gtctgtgttt ttaaaaccac taaagtgcaa gaattacatt 60

<210> 31
 <211> 1416
 <212> DNA
 <213> Homo sapiens

<220>
 <221> Modified base
 <222> 1 ... 1416
 <223> n = a, c, g, or t

<300>
 <308> M55914

<400> 31
 aggaattccg gaattccggga attccgatgg atggaacaga aaataaatct aagtttggtg 60
 cgaacgccat tctgggggtg tcccttgccg tctgcaaaagc tgggtccggt gagaaggggg 120
 tccctctgac cgccacatcg cgtacttgcc tggcaacttc gaagtcatcc tgccagtcoc 180
 ggcgttcaag tgtcatcatc aatggcggtt ctcatgctgy caacaagctg gccatgcaga 240
 gtctgtcctc ccagtcggtg cagcaaaactc agggaaagcca tgcgcgattg gcagcagaagt 300
 ttaccacaac ctgaagaatg tcatcaaggga gaaatatggg aaagatgcca ccaatgtggg 360
 gatttgcgcg ggtttgctcc caacatcctg gagaataaag aaggcctgga gctgctgaag 420
 actgctattg gaaagcctgg cctacactgt aaagttggtc atggcatgga cgtagcggcc 480
 tccagattct tcaaggtcagg gaactatgac ctggaactca agtctcccca tgaccocgcc 540
 aggtacatct cgcctgacca gctggctgac ctgtacaagt ccttcataca ggaactacca 600
 gtgtgtctta tcgaagatcc ctttgaccag gatgactggg gagcttcaga agttcacagc 660
 cagtgacagg atccaggtag tggggggatg actcacagct accaacccaa aggaggtatcg 720
 caaggcgtga acgagaagtc ctgcaactgc ctctgtctca aagtcaccca gattggctcc 780
 gtgacagagt ctcttcaggg gtgcaagctg gccacggcca atggttgggg cgtcatggtg 840
 tctcatcgtt cggggggagac tgaagatacc ttcatcgctg acctggtgtg ggggctgtgc 900
 actggggcag atcaagactg gtgcccttgg ccatgacagc gcttgggcaa gtacaaccag 960

ctcctcagaa	ttgaagagga	gctggggcagc	aaggctaagt	tggccggcag	gaacttcaga	1020
aaccctctgg	ccaagtaagc	tgtgggcagg	caagccttcg	gtcacctgtt	ggctacagac	1080
ccctccctctg	gtgtcagctc	aggcagctcg	aggcccccga	ccaacacttg	cagggggtccc	1140
tgtagtctag	cgccaccgcg	cgtggagttc	gtaccgcttc	cttagaactc	tacagaagcc	1200
aagctccctg	gaagccctgt	tggcagctct	agctttgcag	ttgtgttaatt	ggcccaaatc	1260
attgtttttc	tgccttactt	ttccaccaag	tgtctagagt	catgtgagcc	ttngtgcatac	1320
tcgggggtg	ccacaggcta	gatccccgtg	ggtttttgtc	tcaaaaataa	aagcctcagt	1380
gaccatgaa	aaaaaaaaa	gaattccgga	attccg			1416

<210> 32
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> M55914

<400> 32
 gtaccgcttc cttagaactc tacagaagcc aagctccctg gaagcctgt tggcagctct 60

<210> 33
 <211> 2517
 <212> DNA
 <213> Homo sapiens

<300>
 <308> M96577

<400> 33
 ggaattccgt ggcggggact ttgcaggcag cggcggcggg gggcggagcg ggatcgagcc 60
 ctgcgcgagc cctgcgcgca tgggcgcgcg ccgcgcgcgc cgctgtcac ccgggocgcg 120
 cgggocgtga gcgtcatggc cttggcgcgg gccctctcgg cgggcccactg cgcgcgcgcg 180
 ctggaggccc tgcctggggc cggcgcgctg cggtctgtcg actcctcgca gatcgtcatc 240
 atctccgccc cgcaggacgc cagcgcgccg ccggtcccca cgggcccgcg ggcccccgcg 300
 cgggcccctt cgcagccctga cctgctgtct ttgcacacac cgcaaggccc ccggcccaca 360
 cccagtgcgc cgcggcccgc gctcggccgc ccgcggtga agcggaggct ggacctggaa 420
 actgaccatc agtaccctggc cgagagcagt gggcagctc gggcgagag ccgccatcca 480
 gaaaaaggtg tgaattcccc gggggagaag tcacgctatg agacctcaat gaattctgac 540
 accaagcctt tctggagct gctgagccac tcggctgaag gtgtcgtcga cctgaactcg 600
 gctgcggagg tgcctgaagt gcagaagcgg gcacatctat acatcaccaa cgtccttgag 660
 ggcatccagc tcattgccaa gaagtcacaa aaccacatcc agtggctggg cagccacacc 720
 acagtggggc tgggggacg ccttgagggg ttgaccaggg acctccgaca gctcgaggag 780
 agcagcagc agctggacca cctgatgaat atctgtacta cgcagctcgc cctcctctcc 840
 gaggacactg acagccagcg cctggcctac gtgacgtgtc aggaacctcg tagcattgca 900
 gacctctgag acagataggt tatggtgatc aaagcccctc ctgagaccga gctccaagcc 960
 gtggaactctt cgcagaactt tcagatctcc cttaagagca acaaggcccc gatcgatggt 1020
 ttcctgtgcc ctgaggagac cgtaggtggg atcagccctt ggaagacccc atccagggag 1080
 gtccattctg aggaggagaa caggggccact gactctgcca ccatagtgtc accacaccca 1140
 tcatctcccc cctcatccct caccacagat ccagccactg ctctactcag ctgaggacaa 1200
 gaaccgctgt tgtcccgat gggcagcctg cggctcccg ttgacgagga ccgctgtacc 1260
 ccgctggtgt cggcgcagct gctcctggag catgtgcggg aggaactctc cggcctctcc 1320
 cctgaggagt tcatcagctc ttccccaccc caagaggccc tcgaactacca cttcgccctc 1380
 gaggaggggc agggcactga agacctcttc gactgtgaet ttggggacct caaccccctg 1440
 gattctcgac agggcttggg gggacaaggg ttctcagagt agctcaactt gtctctcgag 1500
 ccttgagacc cctgtctcct ggcgctctcc ccagcctgtt tggaaacatt taattatac 1560
 cctctctctc tgtctccaga agctcttagc tctgggggtc tctgtgtgtg ggctaccgtt aggaggtcta 1620
 gcaagccagg aagggaaggg gtctgtgtgg tctgtatgtg catgcagcct acacccacac 1680
 tgtgtaccg ggggtgaatg tgtgtgagca tgtgtgtgtg catgtaccgg ggaatgaagg 1740
 tgacatata cctctgtgtg tgcactgcag tgcagcccca gtgtgtccac attgtgtgtc 1800
 atgagtcact cctctgcgct gggggggctc taactgcact ttcggccctt ttgctcgtgg 1860

ggtccaccaa	ggcccgaggc	agtgcctgct	cccagaatct	ggtgctctga	ccaggccagg	1920
tggggaggct	ttggctggct	ggggctgtag	gacggtgaga	gcacttctgt	cttaaagggt	1980
ttttctgatt	gaagctttta	tggagcggtta	tttatttatt	gaggcctctt	ttggtgagcct	2040
ggggatcag	caaaaaggga	ggaggggtgt	ggggttgata	ccccaaactcc	ctctaccctt	2100
gagcaaggcg	aggggtccct	gagctgttct	tctgcccat	actgaaggaa	ctgagggcctg	2160
ggtgatttat	ttattgggaa	agtgagggag	ggagacagag	ccatgggtgg		2220
tcagatggtg	gggtggggcc	tctccagggg	gccagttcag	ggcccagctg	ccccccaggga	2280

tggatatgag	atggggagag	tgaatggggg	acottcactg	atgtggggag	gagggggggt	2340
gaagcgctcc	cccagccacg	accctgtggt	ccctcctgca	gtgtctgaag	cgctcgcttc	2400
cccactgctc	tgccccacc	tcacaatctgc	actttgattt	gcttcttaac	agctctgttc	2460
cctcctgctt	tggttttaat	aaatatatttg	atgacgttaa	aaaaaggaat	tcgatat	2517

<210> 34
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> M96577

<400> 34
 gtaggcagggt gagagcactt ctgtcttaaa ggttttttct gattgaagct ttaatggagc 60

<210> 35
 <211> 4437
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000057

<400> 35
 gcgcggcggc cgtggttgcg gcgcgggaag ttggtacct ggttcgcgc gctaggagtc 60
 tgcgtgcgag gattattggct gctgttctct aaaaataatct acaggagcaa ctagaacgtc 120
 actcagccag aacacttaat aataaattaa gtctttcaaa accaaaattt tcagggtttca 180
 cttttaaaaa gaaaacatct tcagataaca atgtatctgt aactaatgtg tcagtagcaa 240
 aaacacctgt attaaagaaat aaagatgtta atgttaccga agacttttcc ttcagtgaac 300
 ctctacccaa caccacaaat cagcaaaagg tcaggactct ctttaaaaaat gctccagcag 360
 gacaggaaac acagagaggt ggtacaaaat cttatttgcc actatttctg cagactccga 420
 aggaagtgtg atgcactacc caaaacacac caactgttaa gaaatccggg gatactgctc 480
 tcaagaaatt agaatttagt tcttaccag attctttaag taccatcaat gattgggatg 540
 atatggatga ctttgatact tctgagactt gctcagaaat caaaaagggt taagagaaac ttttttaagg 600
 accttgttaag agtaagcaact acagtaaaaga ctgatttgcc tccacctcc tctgaagcgc 720
 cacagcttta agcaaaacac gaaacagaag atgactcaga atggttaagc agcgatgtga 780
 agcacaataga tttagactgag gaacagaagg atgtctgaag tgcatataaa tgaagatgtc caggaaagtg 840
 ttgcatcga tgatggcccc attgctgaag gaagtgaaga gagataatag cgaaaagaag aagaatttgg 900
 actctctgaa aactcatttg ttccaattga actgagaag ttccaatgtat gatgatgatt 960
 aagaagctga attacattca cctctccag aagaaattat ttctgctct tcttctctct 1020
 atgatcggga ttttgttcca aaggaacctg acacatctga cagaaaagg gatgttctta 1080
 caaaatgcct tagtactgtta tcaaaacctg agaaaatgag tatgcaggag ctgaatccag 1140
 gcacatcaaa agactcttttg gctagacaga taagttttca gcacagcgtt attctatgtga 1200
 aaaccagcac agactgtgac attgatacta ttctgatga taaactgaaa cttttggatt 1260
 tggagccat ctgtaaatta cagcgggaaca cttcttggtc cattgtggag atacaggcct gattcaactg 1320
 gtgggaacga actgcttcag ctctctggct cagggagatt tatgaaggag ttaaaatttt 1440
 ttaataaaag tgatgccagt tctctgcccc agggactgtt accctaggaa 1500
 atggccctat ggaggggtgat tcttctccct aggaagaatc tttttgaaag gcctttattc aataccattt 1560
 cacacacttc ccaaatctct gttctccctg agcnaactgg caagaacacc aaaaaaaatg 1620
 agacaggatt cttctgccacc ctttggatgt caagcactgc tgtgaaagat cagaataaac 1680
 tcacagaagt ctttggatgt tttccaggga aatgttctca

atactgcttc	aataaatgac	ttagaaagag	aaacccaacc	ttcctatgat	attgataatt	1740
ttgacataga	tgactttgat	gatgatgatg	actgggaaga	cataatgcat	aattttagcag	1800
ccagcaaatc	ttccacagct	gcctatcaac	ccatcaagga	aggtcggcca	attaaatcag	1860
tatcagaaga	acttttccoa	gccaagacag	actgtcttcc	agtgctcatct	actgctcaaa	1920
atataaaact	ctcagagatca	atccagaatt	atactgacaa	gtcagcacaa	aattttagcat	1980
ccagaaactc	gaacatcgag	cgtttccaaa	gtctctagttt	tcctcataca	aaggaaatga	2040
tgaagatttt	tcataaaaaa	tttgccctgc	ataattttag	aactaatcag	ctagaggcga	2100
ccaatgctgc	actgcttggt	gaagactgtt	ttatctcgat	gcgcactgga	ggttggaaga	2160
gtttgtgtta	ccagctccct	gcctgtgttt	ctcctgggtg	cactgtgtgc	acttctccct	2220
tgagatcaact	tatogtagat	caagtccaaa	agctgaactc	cttggtatatt	ccagctccat	2280
atctgcacagg	tgataaagct	gactcagaag	ctacaaatat	ttacctccag	ttatcaaaaa	2340
aagacccaat	cataaaactt	ctatatgtca	ctccagaaaa	gatctgtgca	agtaacagac	2400
tcattttctac	tcctggagaat	ctctatgaga	ggaagctctt	ggcacgtttt	gtattgatgt	2460
aagcacattg	tgctcagtcag	tggggacatg	attttctgca	agattacaaa	agaatgaata	2520
tgcttcgcga	ggacttttct	tctgttccgg	tgatggctct	tacggccaca	gctaattccca	2580
gggtacagaa	ggacatctct	actcagctga	agattctcag	aacctcagtg	tttagcatga	2640
gccttaacag	ccataactctg	aaataactatg	tattaccgaa	aaagccataa	aagggtggcat	2700
ttgattggct	agaatggatc	agaaagcacc	acccatataga	ttcagggata	atttactgcc	2760
ctctccaggcg	agaatgtgac	accatggctg	acacgttaca	gagagatggg	ctcgcctgctc	2820
ttgcttacca	tgctggcctc	agtgatctct	ccagagatga	agtgccagcag	aagtggaata	2880
atcaggtagg	ctgtcaggtt	atctgtgcta	caattgcatg	tggaatgggg	attgacaacac	2940
cggaactggc	atttgtgatt	catgcattct	tccttaaatc	tgtggagggt	tactaccaag	3000
aatctggcag	agctgggaaga	gatggggaaa	tatctcaact	cctgcttttc	tatacctatc	3060
atgatgtgac	cagactgaaa	agacttataa	tgatggaaaa	agatggaaac	catcatacaa	3120
gagaaactca	cttcaaatat	ttgtatagca	tggtacatta	ctgtgaaaa	ataacggaat	3180
gcaggagaat	acagcttttg	gctacttttg	gtgaaaatgg	atttaactct	gatttttgta	3240
agaaacaccc	agatgtttct	tgtgataaatt	gctgtaaaac	aaaggattat	aaaaccaagag	3300
atgtgactga	cgatgtgaaa	agatttgtaa	gatttgttca	agaacatagt	tcatacaag	3360
gaatgagaaa	tataaaacat	gtaggctcct	ctggaagatt	tactatgaat	atgctggctg	3420
acattttctt	ggggagtgaag	agtgcaaaaa	tcacgtcagg	tataatttga	aaaggatctg	3480
cttatttcacg	acacaaatgc	gaagaacttt	ttaaaaagct	gatacttgac	aagatttttg	3540
atgaaagactt	atataatcaat	gcacatgacc	aggcgatgcg	ttatgtgatg	ctcggaaata	3600
aagcccaaac	tgtactaaat	ggcaatttaa	aggtagactt	tatggaaaca	gaaatttcca	3660
gcagtggtgaa	aaaaacaaaa	gcgttagtag	caaaaagtgc	tcagagggaa	gagatgggta	3720
aaaaatgtct	tgagagaactt	acagaagtct	gcaaatctct	ggggaaagtt	tttggttgct	3780
attacttcaa	tatttttaat	accgtcaact	tcaagaagct	tgacagatct	ttatctctctg	3840
ctcagtaggt	tttgcttcaa	attgatgggt	ttaactgaaga	caaaactggaa	aaatatgggt	3900
cggaagtgtc	ttcagtagta	cagaaatact	ctgaattggac	atcgccagct	gaagacagtt	3960
ccccagggat	aagctgtctc	agcagcagag	gccccgggaag	aagtgccgct	gaggagcgtt	4020
acgaggaat	accogtatct	tcocactact	ttgcaagtaa	aaccagaaat	gaaaaggaga	4080
ggaaaaagat	gccagctctc	caaaaggtcta	agaggagaaa	aactgcttcc	agtggttcca	4140
ggcacaaggg	ggggtctgcc	acatgtagaa	agatatcttc	caaaacgaaa	tcctccagca	4200
tcattggatc	ccagtcaagc	tcacataact	ctcaagcgac	atcaggagcc	aatagcaaat	4260
tggggattat	ggctccacag	aaagcotataa	atagaccggt	tttaagcct	tcatatgcat	4320
ttcctaaca	acogaaatct	aatgtacata	gaccctcttt	cttgtttgtc	agcatctgac	4380
catctgtgac	tataaagctg	ttattcttgt	tataccaaaa	aaaaaaaaaa	aaaaaa	4437

<210> 36
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000057

<400> 36
 taagccttca tatgcattct catacaacc gaattctcaat gtacatagac cctctttctt 60

<210> 37
 <211> 2016
 <212> DNA

<213> Homo sapiens

<300>

<308> NM_000060

<400> 37

gccaagctgga gcgttttcgg ggctgtaaa ggagaatggc gcatgcgcat attcagggcg 60
 gaaggcgcgcg taagagcaga tttgtggctc gcattatgtc tggagccaga agtaagcttg 120
 cctcttttctc ctgcggctgt tacgtgggtg cctctgggagc ccacaccggg gaggaagagcg 180
 tggctgacca tcacgagcgt gaattattatg tggctgcccgt gtatgagcat ccatccatcc 240
 tgagtctgaa ccctctggct ctcatcagcc gccaaagaggc cttggagctc atgaaccaga 300

acottgacat ctatgaacag caagtgatga ctgcagccca aaaggatgta cagattatag 360
 tgtttccaga agatggcatt catggattca actttacaag aacatccatt tatccatttt 420
 tggacttcat gccgtctccc caggtgggta ggtggaaccc atgcctggag cctcaccgct 480
 tcaatgacac agaggtgtctc cagcgccctga gttgtatggc catcagggga gatattgtct 540
 tgggtggcaca tcttgggaca aaggagcctt gtcatagcag tgaccacaag tgcccaaaag 600
 atggggagata ccagtccaac acaaatgtcg tgttcagcaa taatggaacc cttgttgacc 660
 gctaccgtaa acacaacctc tactttgagg cagcattcga tgttcctctt aaagtggatc 720
 tcatcacctt tgataccccc tttgtctgga ggtttggcat cttcacatgc tttgatata 780
 tgttctttga ccttgccatc agagtccctca gagactacaa ggtgaagcat gttgtgtacc 840
 caactgcctg gatgaaccag ctcccactct tggcagcaat tgagattcag aaagcttttg 900
 ctgttgccct tggcatccaac gtcttgccag ctaatgtcca ccaccagct ctggggatga 960
 caggaaagtg catacacacc cctctggagt ccttttggtta ccattgacat gaaaatccca 1020
 aaagtccact tataattgccc caggtggcca aaaatccagt ggttctcatt ggtgcagaga 1080
 atgcaacagg tgaacaggac ccattcccata gtaagttttt aaaaattttg tcaggcgatc 1140
 cgtactgtga gaaggatcgt caggaaagtc actgtgatga ggccaccaag tggaaactga 1200
 atgtctctcc cacatttcaac tctgagatga tgtatgacaa tttcacccctg gtccctgtct 1260
 ggggaagaa aggcctatctc cagctctggt ccaatggcct ctgctgttat ttacttttacg 1320
 agaggccacac cttatccaaa gagctgtatg cctctgggggt ctttgatggg cttcacacag 1380
 tacatggcac ttactacatc caagtgtgtg cctctgggcag gtgtgggggt cttggcttcg 1440
 acacctcgcg acaggaatc acagaggcca cgggatatt tgagtctcac ctgtggggca 1500
 acttcagtac ttctctatc ttctctttgt ttctgacctc agggatgacc ctagaagtcc 1560
 ctgaccagct tggctgggag aatgaccact atttctctg gaaaagttag ctgtctctctg 1620
 ggctgggtgac ggcgctctc tatggcgctg tgtatgagag ggactaggaa aagtgtgtgg 1680
 tctgtggggc ggaactctggc catcatgttg acagccttg acttcacag gctacaagcc 1740
 ctgggaccat ctttctgcct taagggcagg agccacttc tgtggcaccg gattccacc 1800
 tgggaactgt ggaaaaagta ggagaggcag attcctctag tgtcttctc ttaaacctca 1860
 atcatcgaga cattaggggg tattttctgt tcacatttat ctttttcaag ccacattctc 1920
 ctctacaaca tctctcagta tgcgatttgt ctcaagctaa aacaaaaata aatgtcagtt 1980
 tatattttac acatccaaaa aaaaaaaaaa aaaaaa 2016

<210> 38

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000060

<400> 38

tcctctaaca aatctctcag tatgcgattg gtctcaagct aaaacaaaaa taaatgtcag 60

<210> 39

<211> 811

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000269

```

<400> 39
gcagaagcgt tccgtgcgtg caagtgcgtg gaaccacgtg ggtccccggc gcgtttcggg 60
tgctggcggc tgcagccgga gttcaaacct aagcagctgg aaggaacctt ggccaactgt 120
gagcgtacct tcattgcat caaacagat ggggtccagc ggggtcttgt gggagagatt 180
atcaagcgtt ttgagcagaa aggattccgc cttgttggtc tgaatttcatt gcaagcttcc 240
gaagatcttc tcaaggaa caacgttgac ctgaaggacc gtccattctt tgccggcctg 300
tgaaataaca tgactcagg gccggttagt gccattgtct gggaggggct gaattgtggt 360
aagacggggc gagtcatgct cggggagacc aacctgcag actccaagcc ttgggaacct 420
ctgggagact tctgcataca agttggcagg aacattatac atggcagtag ttctgtggag 480
agtcagcaga aggagatcgg cttgtggttt caccctgagg aactggtaga ttacacagagc 540
tgtgtcaga actggtatcta tgaatgacag gagggcagac cactattgctt ttacatcca 600
attccatcata atttggaggg aagctcttgg agctgtgagt tctccctgta cagtgttacc 720
atccccagcc atctgattaa aatgcttctt cccagcatag gattcattga gtgtgttact 780
tcatatggtt gcattgcttt tttttcttcc t 811

<210> 40
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000269

<400> 40
gtctgaaatt catgcaagct tccgaagatc ttctcaagga acactacggt gacctgaagg 60

<210> 41
<211> 2338
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000291

<400> 41
agcgacagtc ggcagtcggc tccctcgttg accgaatcac cgacctctct cccagctgt 60
atttccaaaa tgcgcctttc taacaagctg acgctggaca agctggacgt taaagggaag 120
cgggtcgttta tgagagtcga cttcaatggt cctatgaaga acaaccagat aacaaaacac 180
cagagattata aggtcgtctg cccaagcatc aaattctgct tggacaatgg agccaagtgc 240
tgatgcctta tgagccacct aggcgggctc gatggtgtgc ccatgcctga caagtactcc 300
ttagagccag ttgctgtaga actcaaatct ctgctgggca aggatgttct gttcttgaag 360
gactgtgtag gccccagaagt ggagaaagcc tgtgccaaac cagctgctgg gtctgtcatc 420
ctgctggaga acctcgcgtt tcatgtggag gaagaaagga agggaaaaga tgcttctcgg 480
aaccaggtta aagccgagcc agccaaaaa gaagctttcc gagcttcaat ttccaagtca 540
gggagtgtct atgtcaatga tgcttttggc actgctcaca gagcccacag ctccatggt 600
ggagtcaatc tgccacagaa ggctggtggg tttttgatga agaaggagct gaactacttt 660
gcaaaaggct tggagagccc agagcgaccc ttctgggcca tctcggggcg agctaagaat 720
gacagcaaga tccagctcat caataatatg ctggacaaag tcaatgagat gattattggt 780
ggtggaatgy cttttacctt ccttaaggtg ctcaacaaca tggagattgg cacttctctg 840
tttgatgaag agggagccaa gattgtcaaa gacctaatgt cgaactaatg gaagattggt 900
tggaagatta ccttgcctgt tgactttgtc actgctgaca agtttgatga gaagtccaag 960
actggccaag ggtgtattga ttctggcata cctgctggct ggatgggctt ggaattgtgt 1020
cctgaaagca gcaagaagta tgctgaggct gtcactcggg ctaagcagat tggttgaat 1080
ggtctcgtgg ggtgtattga atgggaagct ttgtcccggg ggaactaaag tctcatggat 1140
gagtggttga aagccacttc taggggctgc atcaccatca taggtggttg agacactgcc 1200
acttctgtgt ccaaatggaa cagggaggat aaagtcaagg atgtgagcac tgggggtggt 1260
gccagtttgg agctcctgga agttaaagtc ctctctgggg ttgagtctct cagcaattatt 1320
tagtaacttt ctgcctttta gttcctgtgc acagccccta agtcaactta gcattttctg 1380
catctccact tggcattagc taaaacacct cgtctcaaga ttcagctagt ggtccaagga 1440
tgcaagtcca ggaacccctta aacagttgca cagcatctca gctcatcttc actgcacctc 1500

```

```

ggatttgcac acattcttca agatccatt tgaatttttt agtgactaaa ccattgtgca 1560
ttctagagtg catatatctta tattttgcct gttaaaaaga aagtgcagcag tgttagctta 1620
gttctctttt gatgtagggtt attatgatta gctttgtcac tgtttcacta ctacagcatgg 1680
aaacaagatg aaattccatt tgtaggtagt gagacaaaat tgatgatcca ttaagtaaac 1740
aataaaagtg tccattgaaa ccgtgatttt tttttttttc ctgtcactat ttgttaggaa 1800
gggtgagaat agaattctga ggaacggatc agatgtctat attgtcgaat gcaagaagtg 1860
gggcagcagc agtgagagaga tgggacaatt agataaatgt ccattcttta tcaagggcct 1920

```

```

actttatggc agacattgtg ctatgtcctt tattctaact tttattttta tcagttacac 1980
atgatcataa tttaaaaagt caaggcttat aacaaaaaag cccagcccca ttctcccat 2040
tcaagattcc cactcccacg aggtgaccac tttaactctt tgagtttttc aggtataatac 2100
ctccattgtt ctaagtaata tgcttatatt gttaacttcc ttttttttta ttttttaag 2160
aaatctattt cataccatgg aggaaggctc tgttccacat atatttccac ttcttcatte 2220
tctcgttata gttttgtcac aattatagat tagatcaaaa gtctacataa ctaatacagc 2280
tgagctatgt agtatgctat gattaaattt acctatgtaa aaaaaaaaaa aaaaaaaa 2338

```

```

<210> 42
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_000291

```

```

<400> 42
acttagcatt ttctgcactc ccacttgcca ttactaaaaa ctttccatgt caagattcag 60

```

```

<210> 43
<211> 787
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_000363

```

```

<400> 43
ctgaaggtaa cccggggcgc cccctcactg accctccaaa cgccctctgc ctgcgccctgc 60
ctctgcgcat tcccgccctg agtctcagca tggcgggatg gagcagcagat cgggctaggg 120
aacctcgccc tgcaccagcc ccaatcagac gccgctctct caactaccgc gcttatgcc 180
cggagccgca cgcacaagaa aaatctaaga tctccgctct gaaaaattg cagctgaaga 240
ctctgctgct gcagattgca aagcaagagc tggagcgaga ggcggaggag cggcgaggag 300
agaaggggcg cgtctcgagc acccgctgcc agccgctgga gttgacccgg ctgggcttcg 360
cggagctgca ggaactgtgc cgacagctcc acgcccgtgt ggacaagggt gatgaagaga 420
gatacagcat agaggcaaaa gtcaccaaga acatcacgga gattgcagat ctgactcaga 480
agatctttga ccttcgaggg aagtttaagc ggcccaccct cgggagagtg aggatctctg 540
cagatgccat gatgcaggcg ctgctggggg cccgggctaa ggagtcacct gacctgcggg 600
cccacctcaa gcaggtgaag aaggaggaca ccgagaagga aaaccgggag gtgggagact 660
ggcggaagaa catcgatgca ctgagtgga tggagggccg caagaaaaag tttgagagct 720
gagccttctc gccctactgc cctgccctga ggagggccac tgaggaataa agcttctctc 780
tgagctg 787

```

```

<210> 44
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_000363

```

```

<400> 44
tgtggacaag gtggatgaag agagatacga catagaggca aaagtaccca agaacatcac 60

```

<210> 45
<211> 1263
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000365

<400> 45
ggcaccgagac ctccagcgcc tcggctccag ccccatggcg cccctccagg agttcttcgt 60
tggggggaac tgggaagatga acggcgcgaa gcagagtcgt ggggagctca tcggcactct 120
gaacgcggcc aaggtgccgg ccgacaccga ggtgggttgt gctcccctca ctgcctatat 180
cgactctgcc cggcagaagc tagatcccaa gatgtctgtg gctggcaga actgctacaa 240
agtactaat ggggcttcta ctggggagat cagccctggc atgatcaaa actgcggagc 300
cacgtgggtg gtcctggggc actcagagag aaggcatgtc ttggggagat cagatgagct 360
tatggggcag aaagtggccc atgctctggc agagggactc ggagtaatcg cctgcattgg 420
ggagaagcta gatgaaaggg aagctggcat cactgagaag gttgttttcg agcagacaaa 480
ggcatcgca gataacgtga aggactggag caaggctgtc ctggcctatg agcctgtgtg 540
ggccatttgt actggcaga ctgcaacacc ccaacaggcc cagggaagta acgagaagct 600
ccgagatagg ctgaagtcca acgtctctga tgcggtggct cagagcacc gtatcattta 660
tggaggctct gtagctgggg caacctgcaa ggagctggcc agccagcctg atgtggatgg 720
cttctgtgtg ggtgggtcct cctcaagcc cgaattcgtg gacatcatca atgccaaaca 780
atgagcccca tccactcttc ctacccttcc tgccaagcca gggactaagc agcccagaag 840
cccagtaact gccctttccc tgcataatgt tctgatgggt tcatctgtct ctctcgtgtg 900
cctcatccaa actgtatctt cctttactgt ttatatcttc accctgtaat ggttgggacc 960
agggccaatcc ctctccaact tactataatg gttggaacta aacgtcacca aggtggcttc 1020
tccttggctg agagatggaa ggcgtgggtg gatttgctcc agggcctagt 1080
gagggcagaa gagaaacat cctctccctt ctacacccgt gaggccaaga tcccctcaga 1140
agggcagagt gctgccctct cccatggtgc ccgtgcctct gtctgtgtga tctgaaacc 1200
accatgtgagg gaataaacct ggcactagga aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1260
aaa 1263

<210> 46
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000365

<400> 46
tatcttcacc ctgtaatggt tgggaccagg ccaatccctt ctccacttac tataatggtt 60

<210> 47
<211> 1616
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000582

<400> 47
ctccctgtgt tgggtggagg tgtctgcagc agcattttaa ttctgggagg gcttgggtgt 60
cagcagcagc aggagaggcc agagcacagc atcgtcggga ccagactcgt ctcaggccag 120
ttgcagcctt ctacgacaaa cggcagccaa ggaataactca ctaccatgag aattgcagt 180
atttgccttt gccctctagg catcacctgt gccataccag ttaaacaggc tgactctgga 240
agttctgagg aaaaagcagct ttacaacaaa taaccagatg ctgtggccac atggctaaac 300
cctgacccat ctacagaaca gaatctccta gccccaagaa cccctccagg taagtccaac 360
gaaagccatg aaccacatgga tgatatggat gatgaagat gatgatgata acactgatga ttctcaccag 420
caggactcca ttgactcgaa cgactctgat gatgtagat acactgatga ttctcaccag 480

tctgatgagt	ctcaccattc	tgatgaatct	gatgaactgg	tcactgattt	tcccacggac	540
ctgccacgaa	cogaagtttt	cactccagtt	gtccccacag	tagacacata	tgatggccga	600
ggatgatagt	tggtttatgg	actgaggtca	aaatctaaga	agtttgcgag	acctgacatc	660
cagtaccctg	atgtacacaga	cgaggacatc	acctcacaca	tggaagcga	ggagtgaat	720
ggatgcataca	aggccatccc	cgttgcccag	gacctgaagc	cgcttctcga	ttgggacagc	780
cgtgggaagg	acagttatga	aacgagtcag	ctggatgacc	agagtgtcgt	aaaccacagc	840
cacaagcagt	ccagattata	taagcggaaa	gccaatgatg	agagcaatga	gcattccgat	900
gtgattgata	gtcaggaact	ttccaaagtc	agccgtgaat	tccacagcca	tgaatttcac	960
agccatgaag	atatgctggg	tgtagacccc	aaaagtaagg	aagaagataa	acacctgaaa	1020
tttcgtattt	ctcatgaatt	agatagtgca	tcttctgagg	tcaaataaaa	ggagaaaaaa	1080
tacaatttct	cactttgcat	ttagtcaaaa	gaaaaaatgc	tttatagcaa	aatgaaagag	1140
aacatgaagt	gcttctttct	cagtttattg	gttgaatgtg	tatctatttg	agctcgaaaa	1200
taactaatgt	gtttgataat	tagtttagtt	tgtggtctca	tggaactccc	ctgtaaaacta	1260
aaagcttcag	ggttatgtct	atgttcaatc	tatagaagaa	atgcacacta	tcactgtatt	1320
ttaatatttg	ttattctctc	atgaatagaa	atttatgtag	aagcaaacaa	aatactttta	1380
cccacttaaa	aagagaatat	aacattttat	gtcactataa	tcttttgttt	tttaagttag	1440
tgttatattt	gttgtgatta	tctttttgtg	gtgtgaataa	atcttttatc	ttgaattgaa	1500
taagaatttg	gtggtgtcaa	ttgcttattt	gtttccccc	ggtgtccag	caattaataa	1560
aacataacct	tttttactgc	ctaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaa	1616

<210> 48

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000582

<400> 48

gggtgtgtca attgcttatt tgttttccca cgggtgtcca gcaattaata aaacataacc 60

<210> 49

<211> 1666

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000584

<400> 49

ctccataaagg	cacaaacttt	cagagacagc	agagcacaca	agcttctagg	acaagagcca	60
ggaagaatacc	acccgaagga	accatctcac	tgtgtgtaaa	catgaactcc	aaagctggcc	120
tggtctctctt	ggcagccttc	ctgattttctg	cagctctgtg	tgaagggtga	gttttgccaa	180
ggagtgctaca	agaacttaga	tgctcagtgca	taaaacata	ctccaaacct	ttccacccca	240
aattttatcaa	agaactgaga	gtgattgaga	gtggaccaca	ctgcgcacaac	acagaataa	300
ttgtaaagct	ttctgatgga	agagagctct	gtctggaccc	caaggaaaac	ttgggtgcaga	360
gggtgtggga	gaagtttttg	aaagagggctg	agaattcata	aaaaaattca	ttctctgttg	420
tatccagaaga	tcagtgaaga	tgccagtgaa	acttcaagca	aatctacttc	ataacttcac	480
gtattgtgtg	ggctctgtgt	aggggtgcca	gatgcaatc	aaagattcttg	gttaaatttg	540
aattttcagta	aacaatgaat	agtttttcat	tgtaccatga	aatatccaga	acatacttat	600
atgtaaaagta	ttattttatt	gaatctacaa	aaaacaacaa	ataattttta	aatataagga	660
ttttctcagta	tattgtcagg	gagaataatc	aaatagcaaa	attgaggcca	agggccaaga	720
gaatatccga	acttttaatt	caggaattga	atgggtttgc	tagaatgtga	tatttgaagc	780
atcacataaaa	aatgatggga	caataaattt	tgccataaag	tcaaatttag	ctggaaatcc	840
tggaatttttt	tctgtttaa	ctggcaaccc	tagtctgtca	gcoaggtacc	acaagtctct	900
gttccactgt	gccttgggtt	ctcctttatt	tctaagtga	aaaagtatta	gccaccatct	960
tacctcacag	tgtgtgtgtg	aggacatgtg	gaagcacttt	aaagtttttc	atcataaacat	1020
aaattattttt	caagtgtaac	ttattaacct	attttatttt	tatgtattta	tttaagcatc	1080
aaatatttgt	gcaagaattt	ggaaaaatag	aaagatgaatc	attgattgaa	tagttataaa	1140
gatgttatag	ttaattttatt	ttattttaga	tattaaatga	tgttttatta	gatataaattc	1200
aatcagggtt	tttagatttaa	acaaacaac	aattgggtac	ccagttaaat	tttcatattca	1260

```

gataacaac aaataatttt ttagtataag tacattattg tttatctgaa attttaattg 1320
aactaacaat cctagtttga tactcccagt ctgtgcattg ccagctgtgt tggtagtget 1380
gtgttgaaatt acggaataat gagttagaac tattaaaaa gccaaaactc cacagccaat 1440
attagtaatt tcttgctggt tgaaccttgt ttattatgta caaatagatt cttataatat 1500
tatttaaatg actgcatttt taaatacaag gcttttatatt tttaaactta agatgttttt 1560
atgtgctctc caaatttttt tactgtttc tgattgtatg gaaatataaa agtaaatatg 1620
aaacatttaa aatataattt gttgtcaaag taaaaaaaa aaaaaa 1666

```

<210> 50

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000584

<400> 50

```

tggtagtgt gtgttgaaatt acggaataat gagttagaac tattaaaaa gccaaaactc 60

```

<210> 51

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000599

<400> 51

```

gggaaaaaga gctaggaaga agctgcaaa gctgtgggc ttttccctt ttttgcctc 60
ttttcattac cctcctccg ttttcacct tctccggact tgcgtgaa cctgcgaatt 120
tcgaagagga ggtggcaaa ggggagaaaa gaggtgttag gggttgggtg tttttgtttt 180
ttgttttgtt tttttaattt ctgtatttca acattttctc ccacctctc ggctgcagcc 240
aacgcctctt acctgttctg cggcgccgcg ccgcgtggc agctgagggg tagaagcggg 300
ggtgtatttt agattttaag caaaaaattt aaagataaat ccatttttct ctcccacccc 360
caacgccatc tccactgcac ccgatctcat tattcgggt gttgcttggg ggtgaacaat 420
tttgtggctt tttttccctt ataattctga ccgcctcagg cttgaggggt tctccgcct 480
ccgctcactg cgtgcacctg gcgctgcctt gctcccccac acctgttgca aggttttaat 540
ctttgcaact ggaactctgt cgcaggcacc ccagccctcc acctctctct acatttttgc 600
aagtgtctgg gggagggcac ctgctctacc tgccagaaat tttaaaaaaa aacaaaaaac 660
aaaaaaactc ccggggggccc tcttggcccc ttatccctgt caactctgct ctctctcccc 720
accgccaggt aaaggggggc actaagagaa gatgtgttg ttcacccggg tctcctgtgt 780
gctggcgccc tatggcgggc cggccagag cctgggctcc ttegtgcat gcgagccctg 840
cgacgagaaa gccctctcca tgtccccccc cagccccctg ggctgcgagc tggtaacagg 900
gcccgggtgc ggtgtgtgca tgacctgcgc cctggccgag ggccagtcgt gcgctgtcta 960
caccgagcgc tgcgcacagg ggtgcgctgt cctccccggg caggacgagg agaagccgct 1020
gcacgcctgt ctgcacggcc ggggggtttg cctcaacgaa aagagctacc gcgagcaggt 1080
caagatagag agagactccc gtgagcacga ggagcccacc acctctgaga tggccagga 1140
gacctactcc cccaagatct tccggcccaa acacaccgca atctccgagc tgaaggctga 1200
agcagtgaag aaggacgcga gaaagaagct gaccagctcc aagtttgtcg ggggagccga 1260
gaacactgcc caccoccgga tcatctctgc acctgagatg agacaggagt ctgagcaggg 1320
ccctgcgcgc agacacatgg aggtctccct gcaggagctc aaagccagcc caccgatgtg 1380
gccctgctgt gtgtacctgc ccaatttgga ccgcaaaaga tctacaaga gaaagcagtg 1440
caaaccttcc cgtggcccca agcgtggcat ctgctggtgc gtggacaagt accggatgaa 1500
gctgcacagc atggagtagc ttgacgggga ctttcagtc caacacttcg acgacagcaa 1560
cgttgagtga tgcgtccccc cccaaccttt cctccacccc ctcccacccc cagccccgac 1620
tccagccagc gccctccctc accccaggac gcactctatt tcatctcatt taagggaaaa 1680
atatatatct atctatttga ggaaaaaaa aaaaaaaaaa aa 1722

```

<210> 52

<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_000599

<400> 52
ccaggacgcc actcatttca tctcatttaa gggaaaaata tatatctatc tatttgagga 60

<210> 53
<211> 704
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000735

<400> 53
gcagttactg agaactcata agacgaagct aaaatccctc ttcggatcca cagtcaaccg 60
ccctgaacac atcctgcataa aagcccagag aaaggagcgc catggattac tacagaaaaa 120
atgcagctat ctttctgggc acattgtcgg tgttcttgca tgttctccat tccgctcctg 180
atgtgcagga ttgccagaa tgcacgctac aggaaaaccc attctctccc cagccgggtg 240
cccataact tcagtcgatg ggctgctgct tctctagagc atatcccaat ccactaagg 300
ccaagaagac gatgttgctc caaagaacg tcacctcaga gtccacttgc tgtgtagcta 360
aatcatataa cagggtcaca gtaatggggg gtttcaaat ggagaaccac acggcgtgcc 420
actgcagcac ttgttattat cacaaatctt aaatgtttta ccaagtgctg tcttgaagac 480
tgctgatttt ctggaatgga aaattaagtt gttagtggtt tatggctttg tgagataaaa 540
ctctcctttt ccttaaccata ccaatttgac acgcttcaag gatatactgc agctttactg 600
ccttccctct tatcctacag tacaatcagc agtctagttc ttttcatttg gaatgaatac 660
agcattaaac ttgttccact gcaataaaag ccttttaaat catc 704

<210> 54
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000735

<400> 54
tgagataaaa ctctcctttt ccttaaccata ccaatttgac acgcttcaag gatatactgc 60

<210> 55
<211> 1342
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000799

<400> 55
cccgagagcc gaccggggcc accgcgcccg ctctgctccg acaccgcgcc cctcggacag 60
ccgcctctct ctccaggccc gtggggctgg ccctgcaccg ccgagcttcc cgggatgagg 120
gcgcccggtg tggtcaccgg gcgcgcccca ggtcgtctag ggaccccgcc caggcgcgga 180
gatgggggtg cagcaatgtc ctgcctggct gtggcttctc ctgtccctgc tgtcgtctcc 240
tctggccctc ccagtctctg gcgcgccacc acgcctcatc tgtgacagcc gagtctcgga 300
gaggtacctc ttggaggcca aggaggccga gaatatcacg acggcgtctg ctgaacactg 360
cagottgaat gagaatatca ctgtccacga caccaaagt aatttctatg cctggaagag 420
gatggaggtc gggcagcagg ccgtagaagt ctggcagggc ctggccctgc tgtcggaagc 480
tgctctgggg ggcagggccc tgttggctca cttctccag ccgtgggagc cctcgagct 540
gcatgtggat aaagcgtca gtggccttcg cagcctcacc actctgcttc gggctctgag 600

```

agccacagaag  gaagccatct  cccctccaga  tggggcctca  gctgctccac  tccgaacaat  660
aactgctgac  actttccgca  aactcttcgc  agtctactcc  aatttctccc  ggggaagaat  720
gaagctgtac  acaggggagg  cctgcaggac  aggggacaga  tgaccagggt  tgtccacctg  780
ggcatatcca  ccacctccct  caccacatt  gcttgtgcca  caccctcccc  ccacctccct  840
gaaccccgtc  gaggggctct  cagctcagcg  ccagcgtgtc  ccattggcac  tccagtgtca  900
gcaatgcacat  ctccagggcc  agaggaactg  tccagagagc  aactctgaga  tctaaggatg  960
tcacagggcc  aacttgaggg  ccagagcgag  gaagatttca  gagagcagct  ttaaaactcag  1020
ggacagagcc  atgctgggaa  gacgcctgag  cctactcgcc  accctgcaaa  atttgatgcc  1080
aggacacgct  ttgagggcga  ttaccctgtt  ttgcaccta  ccattcagga  caggatgacc  1140
tggagaactt  aggtggcaag  ctgtgacttc  tcagggtctc  acgggcatgg  gcactccctt  1200
ggtggcaaga  gcccccttga  caccggggtg  gtgggaacca  tgaagacagg  atgggggctg  1260
gcctctggct  ctcatggggt  ccaagttttg  tgtattcttc  aacctattg  acaagaactg  1320
aaaccaccaa  aaaaaaaaaa  aa  1342

```

<210> 56

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000799

<400> 56

```

tcatgggggc  caagttttgt  gtattcttca  acctcattga  caagaactga  aaccacaaaa  60

```

<210> 57

<211> 2722

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000917

<400> 57

```

gagcgggctg  agggtaggaa  gtagccgctc  cgagtggagg  cgactggggg  ctgaagagcg  60
gcgcgccctc  tctgtccact  ttccagggtg  gtgactcctg  aaaattaaat  ctccaagat  120
gatctggtat  atattaatta  taggaattct  gcttccccag  tctttggctc  atccaggctt  180
ttttacttca  attggtcaga  tgactgattt  gatccatact  gagaagatc  tgggtacttc  240
tctgaagat  tatattaagg  cagaagagga  caagtagaaa  caaataaaaa  aatgggcaga  300
gaagttagat  cggctaacta  gtacacgcac  aaaagatcca  gaaggatttg  ttggggatcc  360
agtaaatgca  ttcaaattaa  tgaaacgtct  gaatactgag  tggagttagt  tggagaatct  420
ggtccttaag  gatattgtcag  atggccttat  ctcaaaccta  accattcaga  gccactact  480
ttctaatgat  gaagatcagg  ttggggcagc  caaagctctg  tctactctcc  aggataacta  540
caatttggat  acagatacca  tctcaagggt  taactttcca  ggaagtgaac  acaaatcttt  600
tctaacggct  gagagctgct  ttgagttggg  caaagtggcc  tatacagaag  cagatttata  660
ccatacggaa  ctgtggatgg  aacaagccct  aaggcaactg  gatgaaggcg  agattttctc  720
catagataaa  gtctctgttc  tagattattt  gagctatgog  gtatatcgag  agggagacct  780
ggataaggca  cttttgtcca  caaagaagct  tcttgaacta  gatcctgaac  atcagaagcg  840
taagtgaact  ttaaaatatt  ttgagtatat  aatggctaaa  aaaaaggggg  ttgctgttga  900
tgcttcagat  gaccaaatct  atcagaaaaa  tacaccaagg  ggggagggta  tctcaattgc  960
ttacttgcca  gagagacaga  agtacgaaat  ctgtgcctgt  ggggagggta  tcaaaatgac  1020
ccctcggaga  cagaaaaaac  tcttttgccg  ctaccatgat  ggaaacogta  atcctaaatt  1080
tatcttggct  ccagctaaac  aggaggtatg  atgggacaa  cctcgtatta  ttgccttcca  1140
tgatatatt  tctgatgcag  aaattgaatt  cgtcaaaagc  ctacgcaaac  caagcgtcag  1200
ccgagctaca  atagatgacc  ctgagactgg  aaaaattgac  acagcacagt  acagagtatc  1260
taagagtgcc  tggctctctg  gctatgaaaa  tctgtgggtg  tctcgaatta  atatgagaat  1320
acaagactca  ctgactcag  atgtttccac  agcagaggaa  ttacaggtag  caaattatgy  1380
agttggagga  cagtatgaac  cccattttga  ctttgaccgg  aaagatgagc  cagatgtctt  1440
caaaagactc  gggacagaaa  atagaattgc  tacatggctg  ttttatatga  gtgatgtgtc  1500
tgcaggagga  gccactgttt  tctcctgaag  gtttgcccca  aaaaaggaa  ccaaggaac  1560
tgctgttttc  tggataatac  tgtttgccag  tggagaagga  gattatagta  caccggcatgc  1620

```



```

agcctgtcca  gtgctagtgt  gcaacaaatg  ggtatccaat  aaatggctcc  atgaacgtgg  1680
acaagaattt  cgaagacott  gtacgttgtc  agaattggaa  tgacaaacag  gcttcccttt  1740
ttctcctatt  ttgttactct  tatgtgtctg  atatacacat  ttccatagtc  ttaactttca  1800
ggagtttaca  attgactaac  actccatgat  tgattcagtc  atgaacctca  tcccatgttt  1860
catctgtgga  caattgtcta  ctttgtgggt  tcttttaaaa  gtaacaagaa  atcatcatat  1920
tgcataaaac  taaagaattc  tgttgggtatc  acagaagaca  aggcagagtt  taaagtgagg  1980
taattttata  ttaagaacct  ttttgggttg  ataaaaacat  aatttgagca  tccagtttta  2040
gtatttcaat  acatctcagt  tgggtgggtg  taagctagaa  tgggctgtgt  gataggaaac  2100
aaatgcctta  cagatgtgcc  tagtggttct  gtttacctag  tgtcttaactc  tgttttctgg  2160
atctgaagac  tagtaataaa  ctaggacact  aactgggttc  catgtgattg  ccccttcata  2220
tgatcttcta  agttgatttt  ttctctccca  agtctttttt  aaagaaagta  tactgtattt  2280
taccaccccc  ctctcttttc  ttttagctcc  tctgtggtga  attaaacgta  ctgtgagtaa  2340
aataatttga  tttttttttt  tttttaatg  gaaagtcctg  cataacaaca  ctgggctctc  2400
ttaactaaaa  tgcctcaccc  ttagcctgtt  tttttatccc  ttttttaaaa  tgacagatga  2460
ttttgttcag  gaattttgct  gtttttctta  gtgctaatac  ctgctctctt  atctctgcta  2520
cagcagggtg  gtaatatgtg  cattctgatt  aaataactgt  ccttaggaga  ctggaagttt  2580
aaaaatgtac  aagtcctttc  agtgatgagg  atgtgattt  tttttaaag  tctttttctt  2640
gaaagaccaa  aatgtttgtt  tttttaagat  tctgaaatgt  gttgtgacaa  caatgacctt  2700
ttatgatctt  taaactcttt  tt  2722

```

<210> 58
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000917

<400> 58
tcttactctg ttttctggat ctgaagacta gtaataaact aggcactaa ctgggttcca 60

<210> 59
<211> 3236
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001109

```

<400> 59
gaccgcggcca  tgcgcggcct  cgggctctgg  ctgctggcgg  cgaatgatgt  gcctgcgatt  60
gcccccagacc  ggcctcggtg  cctcatggag  cagtatgagg  tcgtgttgcc  gcgcgctgtg  120
ccaggccccc  gaggcccgcc  agctctgccc  tcccacttgg  gcctgcaccc  agagaggggt  180
agctacgtcc  ttggggccac  agggcacaa  ttccacctcc  acctgcggaa  gaacagggag  240
ctgctgggtt  ccggctacac  agagacctat  acggctgcca  atggctccga  ggtcagggag  300
cagcctcgcc  cggcaggacca  ctgcttatac  cagggccacg  tagaggggta  cccggactca  360
gccgcagacc  tcagcaacctg  tgccggcctc  aggggtttct  tccagggtgg  tcagagacctg  420
cacctgatcg  agccctcggg  tgaagtggtc  gaggcgggac  ggcacgcctg  gtaccaggct  480
gagcacctgc  tgacagcggc  cgggacctgc  ggggtcagcg  acgacagcct  gggcagctcc  540
ctgggacccc  ggacggcagc  cgtcttcagg  cctcgcccg  gggactctct  gccatcccca  600
gagacccgct  actggagact  gtatgtgttc  gtggacaatg  cagagttcca  gatgctgggg  660
agcgaaagc  ccgtgcgtca  tcgggtgctg  gagggtgtga  atcacgtgga  caagctatat  720
cagaaaactca  acttccgtgt  ggtcctgggt  ggctcggaag  tttggaatag  tcaggcacgg  780
ttccacgtca  gcccccagcc  cagtgctaca  ctggagaacc  tctgacctgc  cgaggcacgg  840
caacggacac  gtcggcacct  gcatgacaac  cgaagctcca  tcacgggtgt  cgactctacc  900
gggactactg  tggsgtttgc  cagggtgtcc  gccatgtgct  cccaacagctc  aggggtctgt  960
aacoaggaac  acagcaagaa  ccccggtggc  gtggcctgca  atctggccca  tgagatgggc  1020
cacaacctgg  gcatggacca  tgatgagaac  gtccagggct  gccgctgcca  ggaacgcttc  1080
gaggccggcc  gctgcatcat  ggcaggcagc  attggctcca  gtttcccag  gatgttcagt  1140
gactgagacc  aggcctacct  ggagagcttt  gtggagcggc  cgcagtcggt  gtcgctcgcc  1200
aacgccctgt  acctcagcca  cctggtgggc  ggcccctgt  gtgggaacct  gtttgtggag  1260

```

cgtaggggagc	agtgcgactg	cgcccccccc	gaggactgcc	ggaaccgctg	ctgcgaactct	1320
accacctgcc	agctggctga	ggggggccag	tgtgcgcacg	gtacctgctg	ccaggagctg	1380
aaggtgaagc	cgctgggtga	gctgtgccgt	cccaagaagc	acatgtgtga	cctcgaggag	1440
ttctgtgacg	gcccgcaccc	tgagtgcctg	gaagacgcct	tccaggagaa	cggcacgcgc	1500
tgtccggggg	gctactgtcta	caacggggcc	tgtccacac	tgccccagca	gtgccaggcc	1560
ttctgggggc	cagggtggca	ggctgccgag	gagtcctgct	ttctctatga	catctacca	1620
ggctgcgaag	ccagccggta	cagggtcgac	atgtgtggcg	ttctgcagtg	caagggtggg	1680
cagcagcccc	tggggcgctg	catctgcatt	gtggatgtgt	gccacgcgct	caccacagag	1740
gtggcgactg	cgtatgaacc	agtgcccag	ggcaccgcgt	gtggaccaga	gaaggtttgc	1800
tgaagaaggc	gttgcccagg	cttacacgtt	tacagatcca	gcaactgctc	tgcccagtcg	1860
cacaaccatg	gggtgtgcaa	ccacaagcag	gagtgccact	gccacgcggg	ctggggcccc	1920
cccactgcg	cgaagctgct	gaetgaggtg	cacgcagcgt	ccggagcctt	ccccgtcctc	1980
gtggtggtg	ttctggtgct	cctggcagtt	gtgctggtca	ccctggcagg	catcatcgct	2040
taccgcaaa	ccggagaccg	catcctgagc	aggaacgtgg	ctcccaagac	cacaatgggg	2100
cgcttcgcac	ccctgttcca	ccaggctgcc	agcccgctgc	cgcccaaggg	cggggctcca	2160
gccccatcca	ggggccccca	agagctggct	cccaccaccc	accggggcca	gcccgcgcga	2220
caaccggctc	cctcggtggc	tctgaagagg	ccgccccctg	ctcctccggt	catctgtgct	2280
agcccccctt	tcccagttcc	tgtctacacc	cgccaggcac	caaagcaggt	catcaagcca	2340
acgtttcgac	ccccagtgcc	cccagtcaaa	cccggggctg	gtgcggccaa	ccctggtcca	2400
gctgaggggt	ctgtttggccc	aaaggttgcc	ctgaagcccc	ccatccagag	gaagcaagga	2460
gcccggagctc	ccacagcacc	ctaggggggc	acctgcgcct	gtgtggaaat	ttggagaagt	2520
tgccgcagag	aaegcatgct	ttccagcctt	ccacggtcca	gctagtgcgc	ctcagcccta	2580
gacctcgact	ttgccagctc	agctgctgtt	ctaaccctcag	taatgcattc	acctgagaggt	2640
ctcctgctgt	ccacgccttc	agccaattcc	ttctccccgc	cttgcccacg	tgtagcccca	2700
gctgtctgca	ggcaccaggc	tgggatgagc	tgtgtgcttg	cggtgtgcgt	tgtgtgtacg	2760
gtgtccagg	tggccgctgg	tctcccgcgt	tgttcaggag	gccacatata	cagccccctc	2820
cagccacaacc	tgcacctgct	ctggggccctg	ctgagccggc	tgccttgggc	accggcttcc	2880
aggcagcaca	gacgtggggc	atccccagaa	agactccatc	ccaggaccag	tgtccctccc	2940
gtgctcttgc	agaggggtgc	agtgcagaga	ctgcaccccc	agctcccagc	tccaggtccc	3000
ctgactctgg	ggcctgtttcc	catgggattc	aaagaggaga	gccccagctt	tgtgtgtgtt	3060
taagcttagg	aatgcccttt	atggaaaggg	ctatgtggga	gagtcagcta	tcttgtctcg	3120
ttttcttgag	acctcagatg	tgtgttcagc	agggctgaaa	gcttttattc	tttaataatg	3180
agaaatgtat	attttactaa	taaatatttg	accgagttct	gtagattctt	gttaga	3236

<210> 60

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001109

<400> 60

ctttatggaa agggctatgt ggggagatca gctatcttgt ctggttttct tgagacctca 60

<210> 61

<211> 1449

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001124

<400> 61

ctggatagaa	cagctcaagc	cttgccactt	cgggcttctc	actgcagctg	gcttgggaact	60
tcggagtttt	gccattgcc	gtgggagctc	tgagactttc	tccttcaagt	acttggcaga	120
tcactctctt	agcaggggtc	gcctctcgca	gcccggatga	agctgggttc	cgctgcctcg	180
atgtactctg	gttcgcctgc	cttcttagcg	gtcgacacgc	ctcggttggg	tgtcgcgctg	240
gagttctgaa	agaagttgaa	taagttggct	ctgagtcgtg	ggaaagaggg	actcgggatg	300
tcagcagctg	accaccagg	gctcgctgac	gtgaaggccg	ggcctgccca	gaccttatt	360
cggccccagg	acatgaagg	tgctctcgca	agccccgaag	acagcagctc	ggatgcgcgc	420

cgcatccgag	tcaagcgcta	cgcgcagagc	atgaacaact	tccagggcct	coggagcttt	480
ggctgcogct	tcgggacgtg	cacgggtgcag	aagctggcac	accagatcta	ccagttcaca	540
gataaaggaca	aggacaacgt	cgcccccagg	agcaagatca	gcccccaggg	ctacggcgcc	600
cggcgcggcg	gtctccctgcc	cgaggccggc	ccgggtcgga	ctctgtgtgc	ttctaaagca	660
caagcacaacg	gggctccagc	ccccccgagt	ggaagtgtct	cccactttct	ttaggattta	720
ggcgcccatg	gtacaaggaa	tagtcgcgca	agcatcccg	tgggtcctcc	cgggacgaag	780
gaacttccgca	gcgggtgtggg	gaaccggctc	tgacagccct	gcggagaccc	tgagtcoggg	840
aggcacccgtc	cggcgcgag	ctctggcttt	gcaaggccce	ctcctctcgg	gggcttcgct	900
tecttagcctt	gtctcaggtg	caagtgtccc	agggggcggg	gtgcagaaga	atccagagtgt	960
ttgcagggtc	taaggagagg	agaaaactgag	aaatgaatgc	tgagaccccc	ggagcagggg	1020
tctgagccac	agccgtgtctc	gcccacaaac	tgattttcta	cgccgtgtca	ccccaccagg	1080
gcgcgaagcct	cactattact	tgaactttcc	aaaacctaaa	gaggaaaagt	gcaatgcgtg	1140
ttgtacatac	agaggttaact	atcaatatatt	aagtttgttg	ctgtcaagat	tttttttgta	1200
aottcaaatat	tagagataatt	tttgtacggt	atatatttga	tttaagggat	tttaaaagca	1260
attatattgt	ttccccctat	tttaagacgt	gaatgtctca	gcgaggtgta	aagttgttcg	1320
cgcgctggaa	tgtgagtggt	tttgtgtgca	tgaagagaaa	agactgatta	cctcctgtgt	1380
ggaagaagga	aaacccgagt	ctctgtataa	tctatttaca	taaaatgggt	gatatgcgaa	1440
cagcaaac	1449					

<210> 62

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001124

<400> 62

gaaggaaaca cagagtctct gtataatcta ttacataaa atgggtgata tgcgaacagc 60

<210> 63

<211> 1619

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001168

<400> 63

cgcccaagatt	tgaatcgagg	gaaccttttg	cagagggtggc	ggcggcgcca	tgggtgcccc	60
gacgttgccc	cctgcctggc	agccctttct	caaggaccac	cgcactctta	catccaagaa	120
ctggccctct	ttggagggtc	gcgcctgcac	cccgagcggg	atggccgagg	ctgggttcac	180
ccaactgcgcc	actgagaacg	agccagaact	ggcccagtg	ttcttctgct	tcaaggagct	240
ggaaaggctgg	gagccagatg	acgaccccat	agaggaaact	aaaaagcatt	cgtccgggtg	300
cgtcttccct	tctgtcaaga	agcagtttga	agaattaaac	cttgggtgaat	ttttgaaact	360
ggagacagaaa	agagccaaga	acaaaaattgc	aaaggaaaac	aacaataaga	agaaagaatt	420
tgaggaaact	gcgaagaaag	tgcgcctgtc	catcgagcag	ctggctgcga	tggattgagg	480
cctctggccg	gagctgcctg	gtcccagagt	ggctgcacca	cttccagggt	ttattccctg	540
gtgccaccag	ccttccctgg	ggccctctag	caatgtctta	ggaaaggaga	tcaaacatttt	600
caaatatagat	gttttcaactg	tgtcctctgt	ttgtcttgaa	agtggcacca	gaggtgtctc	660
tgcctgtgca	cggggtgctg	ctggtaacag	tggctgcttc	ttctctctct	ttctctctct	720
gggggtctcat	ttttgtctgt	ttgattcccg	ggcttaccag	gtgagaagtg	agggagggaag	780
aaggcactgtg	cccttttctt	agagctgaca	gctttgttcg	cgtgggcaga	gccttccaca	840
tgaatgtgtg	ctggacctca	tgttgttgag	gctgtcacag	tctcgtagtg	ggaattggca	900
gggtgcctgtt	gaaatctgagc	tgcagggttc	ttctatgtca	caactgtgcc	tctccagagg	960
acagtttttt	tgttgttgtg	tttttttgtt	tttttttttt	ggtagatgca	tgaattgtgt	1020
gtgatgagag	aatggagaca	gagtcctctg	ctcctctaact	gtttaacaaac	atggctttct	1080
tattttgttt	gaattgttaa	ttcacaagat	agcacaacac	acaattaaaa	ctaaagacaa	1140
agccatttcta	agtcatgtgg	gaaacgggtg	gaacttcagg	tggatgagga	gacagaaatg	1200
agtgaaggga	agcgtctggc	agtactcctc	tttgccactg	ctgtgtgatt	atgcagagccc	1260
agtgagccgc	gggggcacatg	ctggccgctc	ctccctcaga	aaaaggcagt	ggcctaatac	1320

```

ctttttaaat gacttggctc gatgctgtgg gggactggct gggctgctgc aggccgtgtg 1380
tctgtcagcc caacctctac atctgtcaac ttctccacac gggggagaga cgcagtcgcg 1440
ccaggtcccc gcttctcttg gaggcagcag ctcccgcagg gctgaagtct ggcgtaaagt 1500
gatggatttg attcgccctc ctccctgtca tagagctgca ggggtgattg ttacagcttc 1560
gtcggaaacc tctggaggtc atctcggtgt ttctgtgaaa ataaaaagcc tgtcatttc 1619

```

<210> 64

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001168

<400> 64

ttcacagaat agcacaact acaattaaaa ctaagcaca agccattcta agtcattggg 60

<210> 65

<211> 1552

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001216

<400> 65

```

gcccgtacac accgtgtgct gggacacccc acagtcagcc gcatggctcc cctgtgcccc 60
agccccctggc tccctctgtt gatcccggcc cctgtccaa gctcactgt gcaactgctg 120
ctgtcacctgc tgcctctgat gctgttccat ccccagaggt tgcctccggt gcaggaggat 180
tcccccttgg gaggaaggctc ttctggggaa gatgaccac tggggcagga ggaactgtccc 240
agtgaagagg attcaccocag agaggaggat ccaccocagg agggagatct accgtggagag 300
gaggtatctac ctggagagga ggaatctacct gaagttaagc ctaaatcaga agaagagggc 360
tccctgaagt tagaggatct acctactgtt gaggctcctg gagatcctca agaaccocag 420
aataatgccc acaggggacaa agaaggggat gaccagagt attggcgcta tggaggcgac 480
cgccccagc cccgggtgtc cccagcctgc gcggggcgct tccagctccc ggtggatctc 540
cgccccagc tccggcctct ctgcccggcc ctgcccctcc tggactcct ggccttccag 600
ctccccgcgc tcccagaact gcgcctgcgc aacaatggcc acagtgctga actgaacctg 660
ctcctctggc tagagatggc tctgggtccc gggcgggagt accgggctcg cagctgcact 720
ctgcactggg gggctgcagg tctccggggc tggagcaca ctgtggaagg ccacctttc 780
ctgcgcgaga tccacgtggt tcacctcagc accgctttg ccaagagtga cgaggctcg 840
ggggcggcgg gaggcctggc cgtgttgccc gcccttctgg agggaggccc ggaagaaaac 900
agtcctctatg agcagtgtgt gtctgccttg gaagaaatcg ctgaggaagg ctccagagact 960
caggtccocag gaetggacat atctgcactc ctgcccctcg acttcccaa ctacttccaa 1020
tatgaggggt ccttgactac accgcccctg gccagggtg tcatctggac tgtgtttaac 1080
cagcagtgta tgcctgagtc taagcagctc cacaccctct ctgacacctc gtggggacct 1140
ggtgactctc ggctacaagct gaacttccga gcgacgcagc ctttgaatgg gcgagtgatt 1200
gagggctctc tccctctggg agtggacagc agtcctcggg ctgctgagcc agtcagctg 1260
aatctctgoc tggctgtggt tgacatccta gccctggttt ttggctcctc tttgtgctc 1320
accagcgtcg cgttctctgt cgagatgaga aggcagcaca gaaggggaac caaagggggt 1380
gtgagctacc gccacgcaga ggtagccgag actggagcct agaggtctga tcttgagaaa 1440
tgtgagaagc cagccagagg catctgagg ggagccggtg agtctcctgt cctgctcatt 1500
atgccacttc cttttaactg ccaagaaatt ttttaaaata aatatttata at 1552

```

<210> 66

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001216

<400> 66
tctgtctctg ctcattatgc cacttccctt taactgccaa gaaatttttt aaaataaata 60

<210> 67
<211> 2653
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001254

<400> 67
gagcgcgctt ggagtttgct gctgccgtgt tgacgtttgt tcaggggctt gtggtgtgta 60
gtccgagagg ctgcgtgtga gagacgtgag aaggatcctg cactgaggag gtggaagaag 120
gaggattgct cgaggaggcc tggggtctgt gaggcagcgg agctgggtga aggtctgggg 180
ttccggcgag gctctgagctg tgctgtcgtc atgcctcaaa cccgatccca ggacacggct 240
acaactcagt ttccaaaag gaaagtgtct cgggcattga acaaagctaa aaactccagt 300
gatgccaacac tagaaccacac aatgtgtcaa accgttaacct gttctcctgt tgaataagcc 360
ctgcctctca gccccaggaa acgtctgggc gatgacaacc tatgcaacac tcccatttta 420
ctctccttgtt ctccaccaaa gcaaggcaag aaagagaatg gtcacctca ctccatata 480
cttaaggagac gaagattgggt attgacaat cagctgacaa ttaagtctcc tagcaaaaga 540
gaactagcca aagttccacca aaacaaaata cttttctcag ttgaaaaag tcaagagatc 600
acaacaaatt ctgagcagag atgtccactg aagaaaagaat ctgcattgtg gagactatcc 660
aagcaagaag gcaattgtcta ccagcaagca aagctgtgtc tgaacacagc tgtcccgat 720
cggctgacct ccagggaag ggagatggat gtcatcagga atttcttgag ggaacacatc 780
tgtgggaaaa aagctggaag cctttacctt tctggtgtct ctggaaactg aaaaactgcc 840
tgcttaagcc ggaattgtca agacctcaag aaggaaactga aaggtcttaa aactatcatg 900
ctgaattgca tgtccttgag gactgccag gctgtattcc cagctattgc tcaggagatt 960
tgtcaggaag aggtatccag gccagctggg aaggacatga tgaggaattt ggaaaaacat 1020
atgactcga agaaaggccc catgatttgt ttggtattgg acgagatgga tcaactggac 1080
agcaaggccc aggatgtatt gtaacacgta ttgtaattgg catggctaa ccaattctcac 1140
ttggtgtgta ttggtattgc taatccctg gatctcaag atagaattct acctaggctt 1200
caagctagag aaaaatgtaa gccacagctg ttgaacttcc cactttatac cagaatacag 1260
atagtacta ttttgcaaga tcgacttaac caggtatcta gagatcaggt tctggacaat 1320
gctgcagttc aattctgtgc ccgcaaatgc tctgctgttt caggagatgt tcgcaaaaga 1380
ctggatgttt gcagagagc tattgaaatt gttagatcag atgtcaaaag ccagactatt 1440
ctcaaacacc tgtctgaatg taatcacct tctgagcctc tgattcccaa gggggtgtgt 1500
cttattacac tatcccaagt catctcagaa gttgatggtt acaggatgac cttgagccaa 1560
gaaggagcac aagattcctt cctctctcag cagaagatct tgggttgtct tttgatgctc 1620
ttgatcagcc agttgaaaat caaagaggtc actctgggga agttatatga agcctacagt 1680
aaagtctgtc gcaaacagca ggtggcggtc gtggaccagt cagagtgttt gtcactttaa 1740
gggctcttgg aagccagggg cattttagga ttaaaagaa acaaggaaac cggtttgaca 1800
aaggtgtttt tcaagattga agagaaagaa atagaacatg ctctgaaaga taaagcttta 1860
attggaataa tcttagctac tggattgcct taaattcttc tcttacaccc caccgaaag 1920

gattacagct gcatttagag agctacagtc ttcatttttag tgctttacac attcgggctc 1980
taaaacaaat atgacctttt ttacttgaag ccaatgaatt ttaactctaa gactctttaa 2040
tattacacaa gaataaatatc tttgggtctt actattttta cccataaaaag tgaccaggtta 2100
gacctttttt aattacattc actactctca ccaattgtct atctctagcc aatgtgtctg 2160
caagtgtaca gatctgtgta gaggaaatgtg tgtatattta cctcttggtt tgctcaaaa 2220
tgagtggtta tttttttgtt tttttttgtt gttgtgtgtg tttttgagc ggcgtccacc 2280
ctgttgacca ggtctggagt caatggcgcg ttctctgtct actacagcac cgcgttccca 2340
ggttggtgca attctcttgc ctacgctcc cagatagctg ggattacagg tgcaccaacc 2400
cgcgccagcc taatttttta attttttagta gagacaggtt ttaccatgt tgccagggct 2460
ggtcttgaa cctgcagacc caagtgtatc gccacctgt gccacctaa gctgtgggat 2520
tataagcgtg agccacacatg ctacgacctt aaggtatttt gttaagaact ttaagtttag 2580
ggtaagaaga atgaaaaatga tccagaaaaa tgcaagcaag tccacatgga gatttggagg 2640
acactggtta aag 2653

<210> 68
<211> 60

```

<212> DNA
<213> Homo sapiens

<300>
<308> NM_001254

<400> 68
caaggaaaccc cgtttgacaa aggtgttttt caagattgaa gagaagaaa tagaacatgc 60

<210> 69
<211> 627
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001323

<400> 69
gcggccgcaa gctcggcact cacggctctg agggctccga cggcactgac ggccatggcg 60
cgttcgaaac tcccgctggc gctgggcctg gccctggctg cattctgctt cctggcgctg 120
ccacgcgcagc cccgggcccg gccgcaggag cgcattgctg gagaactccg ggacctgtcg 180
cccgacgcagc ccgaggtgca gaagcggcgc caggcggcgc tggccagcta caacatggcg 240
agcaacagca tctactactt ccgagacacg cacatcatca aggcgcagag ccagctggtg 300
ccgcgcatac agtacttctt gacgatggag atggggagca cagactgccg caagaccagg 360
gtcaactggag accacgttga cctcaccact tgccccttgg cagcagggcg gcagcaggag 420
aagctgcgct gtgactttga ggtccttgtg gttccctggc agaactcctc tcagctccta 480
aagcacaaat gtgtgcagat gtgataagtc cccgaggcgc aaggccattg ggtttggggc 540
catggtggag ggcacttcag gtcctggggc cgtatctgtc acaataaatg gccagtgctg 600
cttcttgcaa aaaaaaaaaa aaaaaaa 627

<210> 70
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001323

<400> 70
atcaagtact tcttgacgat ggagatgggg agcacagact gccgaagac cagggctcact 60

<210> 71
<211> 1812
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001428

<400> 71
tagctaggca ggaagtcggc gcggggcgcg cggacagtat ctgtgggtac ccggagcagc 60
gagatctcgc cggttttacg ttacacctcg tgctcgcagc accctccgct tctctccta 120
ggcgacgaga ccagtggtct agaagttcac catgtctatt ctcaagatcc atgccaggga 180
gatctttgac tctcgcggga atccactgtg tgaggttgat ctcttcacct caaaaggtct 240
cttcagagct gctgtgccca gtggtgcttc aactcgtatc tatgaggccc tagagctcgc 300
ggacaatgat aagactcgct atatggggaa ggggtgtcca aaggctgttg agcacatcaa 360
taaaactatt gcgcctgccc tgggttagcaa gaacctgaac gtcacagaac aagagaagat 420
tgacaaactg atgatcgaga tggatggaa agaaaaataa tctaagtttg gtgcgaacgc 480
cattctgggg gtgtcccttg ccgtctgcaa agctggtgac gttgagaagg ggttccccct 540
gtaccgcgac atcgctgact tggctggcaa ctctgaagtc atcctgccag tcccgcggtt 600
caatgtcatc aatggcgggt tctatgctgg caacaagctg gccatgcagg agttcatgat 660

```

```

ctccaccatc ggtgcagcaa acctcaggga agccatgcgc attggagcag aggtttacca 720
caacctgaag aatgtcatca agggagaaata tgggaagaat gccaccaatg tgggggatga 780
aggggggttt gctcccaca cctcgtgagaa taagaagagg ctggagctgc tgaagactgc 840
tattgggaaa gctggctaca ctgataaggt ggtcatcggc atggacgtag cggcctccga 900
tcttcttcagg tctgggaagt atgacctgga ctccaagtct cccgatgacc ccagcaggta 960
catctcgccct gaccagctgg ctgacctgta caagtccctt atcaaggact acccagtgtt 1020
gtctatcgaa gatccctttg accaggatga ctggggagct tggcagaagt tcacagccag 1080
tgagggaatc caggtagtgg gggatgatct cacagtgaac aaccacaaga ggatcgccaa 1140
ggccgtgaac gagaaagtct gcaactgcct cctgctcaaa gtcaaccaga ttgctccgt 1200
gaccagctct ctccaggcgt gcaagctggc ccaggccaat ggttggggcg tcatggtgtc 1260
tcattcgttcg ggggagactg aagatacctt catcgctgac ctgggtgtgt ggctgtgtac 1320
tgggcagatc aagactgtgt ccccttgccg atctgagcgc ttggccaagt acaaccagct 1380
cctcagaatt gaagaggagg tgggcagcaa ggctaagttt gccggcagga acctcagaaa 1440
ccccctggcc cactaagctg tgggcaggca agcccttcgg tcacctgttg gctacacaga 1500
ccccctccct cgtgtcagct caggcagctc gaggcccccg accaacactt cgagggttcc 1560
ctgctagtta gcgccccacc gccctggagt tctgaccgct tccttagaac ttctacagaa 1620
gcgaagctcc ctggagccct gttagcagct ctgctcttgc agtctgttaa ttggcccaag 1680
tcattgtttt tctcgctcca ctctccacca agtgtctaga gtcatgtgag cctcgtgtca 1740
tctccggggt ggccacaggc tagatccccg gtgtttttgt gctcaaaaaa aaagcctca 1800
gtgaccatcg ag 1812

```

<210> 72

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001428

<400> 72

agctctagct tttagcagtc tgtaatgggc ccaagtcatt gttttctctg cctcactttc 60

<210> 73

<211> 8368

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001456

<400> 73

```

gcgatccggg gcgccacccg cggctcatcg tcaccgggtg ctctcaggaa cagcagcgca 60
cctgtcgtct cctgcctcgc ctcccgcgcg cttaggtgct tgcgacttta attaaagggc 120
cgtcccctcg ccgaggctgc agcaccgcgc ccccggtctt tcgcgctcca aaatgagtag 180
ctcccactct ccggcgggcc agagcgcgag agggcgcgct cggcgcgcg cgctcgacac 240
cggggagcgc gagatgccg ccaccgagaa ggaacctggc gaggacgcgc cgtggaagaa 300
gatccagcag aacactttca cgcgctgggt caacgagcac ctgaagtgcg tgagcaagcg 360
catcgccaac ctgcagacgc acctgagcga cgggctgcgc ctatccgcgc tgttggaagt 420

```

```

gtcagccag aagaagatgc accgcaagca caaccagcgg cccactttcc gccaaatgca 480
cgttgagaac gtgtcggtgg cgtcgaagtt cctggaacgc gagagcatca aactggtgtc 540
catcgacagc aaggccatcg tggacgggaa cctgaagctg atcctggggc tcactctggac 600
ctgactcctg cactactcca tctccatgcc catgtgggac gaggaggagg atgaggaggc 660
caagaagcag acccccgaag agaggctcct gggctggatc cagaaacaag tggccagact 720
gcccataccc aacttcagcc gggactggca gagcgggcgg gccctgggcg ccttgggtga 780
cagctgtgtc cgggcccgtg gtccctgaact ggaactcttg ggaaccagca agcccgcttac 840
caatgcgcga gaggccatgc agcaggcgga tgaactggct ggaatccccc aggtgatcac 900
cccgaggagg attgtggacc ccaactggga cgaagactct gtactgacct acctgtccca 960
gttccccaa gccaagctga agccaggggc tccttctgcg cccaaactga acccgaagaa 1020
agccctgcc tacggggcag gcatcgagcc caacggcaac acggtgaaga agcgggagaa 1080
gttactgtg gagaccagaa gtgctggcca gggagagggt ctggtgtacg tggaggaccc 1140

```

ggccggacac	caggaggagg	caaaagtgc	cgccaataac	gacaagaacc	gcacctcttc	1200
cgctcgtgtac	gtccccgagg	tgacggggac	tcataaggtt	actgtgctct	tttgtggcca	1260
gcacatcgcc	aaagggccct	tcgaggtgta	cgtggataag	tcacaggggtg	acgccagcaa	1320
agtgacacgc	caaggttcccg	gcctggagcg	cagtggaacac	atcgccaaca	agaccaacta	1380
ctttgagatc	tttacggcag	gagctggcac	gggcgaggtc	gaggtttgtga	tcacaggacc	1440
catggagacg	aaggggacgg	tagagcctca	gctggagggcc	cggggcgcaca	gcacataccg	1500
ctgcagctac	cagcccccaca	tggaggcggt	ccacaccgtg	cacgtcacgt	ttccggcggt	1560
gccatccctc	cgagccccc	acactgtcac	tgttgcccaa	gctgttaacc	cgagtgctctg	1620
ccggcggtgt	ggccggggcc	tcacgcccaa	gggtgtgcgg	gtgaaggaga	cagctgaact	1680
caaggtgtac	acaaaggcgg	ctggcagtg	ggagctgaag	gtcacgtga	agggccccc	1740
gggagaggag	cgctgaaagc	agaaggacct	gggggatggc	gtgtatggct	tcagattatta	1800
ccccatggtc	ctgggaagtc	atatcgtcac	catcacgtgg	ggtggtcaga	acatcgggcg	1860
cagtcocctc	gaagtgaagg	tgggcaccga	gtgtggcaat	cagaaggtac	gggcctgggg	1920
ccctggggtg	gaaggcggtg	tcgttggcaa	gtcagcagac	tttgtgtgtg	aggtcatcg	1980
ggagcagctg	ggcagcgtgg	ggtctcgtgt	ggaaaggcca	tcgcagggta	agatcgaaatg	2040
tgacgacacg	ggcgacggct	ccgtgtgatgt	gcgctactgg	ccgcagagg	ctggcgagta	2100
tgccgttcac	gcgtgtgca	acagcggaaga	catcgcctca	agcccttcca	ttggctgacat	2160
ccgtgacgcg	ccccaggact	tcacaccaga	caggggtgaag	gcacgtgggc	ctggatgtga	2220
gaagacaggt	gtggcgctca	acaagccagc	agacttcaca	gtggatgcca	agcacgggtg	2280
caagggccca	cttcgggttc	aagtccaggga	caatgaaggc	tgccctgtgg	agggcttgg	2340
caaggaccaac	ggcaatggcc	cttacagctg	ctcctacgtg	ccaggaagcg	cggtgaagca	2400
cacagccatg	gtgtcctggg	gaggcgctcag	catccccaac	agcccttcca	gggtgaatgt	2460
gggagctggc	agccacccca	acaaggtcaa	agtatacggc	cccggagtag	ccaagacagg	2520
cgtaacggcc	cacagcccca	cctacttcac	tgtggactgc	gccgaggctg	gccaggggga	2580
cgctcagcat	ggcatcgaat	gtgccctgg	agtggtaggc	cccgcggaag	ctgcacatcga	2640
cttcgcacatc	atccgcaatg	acaatgacac	cttcacgggtc	aagtacacgc	cccggggggc	2700
tggcagctac	accattatgg	tcctcttgc	tgaccaggcc	acgccccaaca	gccccatccg	2760
agtcagggtg	gagccctctc	atgacggccg	taagggtgaag	ggcgaggccg	ctggcctcag	2820
tcgcactggt	ctcgagcttc	gcgaagccac	ccaacttcaca	gtaaatgcca	aaagctgctg	2880
caaaaggcaag	ctggagctgc	agttctcaag	actcaaccaag	ggggagtgcag	tcgcagatgt	2940
ggacatcatt	gaccacatgt	acaacacctc	cacagctcaag	tacacgcctg	ctccagcagg	3000
tcgaataggc	gtcaatgtca	cttatggagg	ggatcccatc	cctaagagacc	ctttctcagt	3060
ggcagtatct	ccaagcctgg	acctcagcaa	gatcaagggt	ttctggcctg	gagagaaggt	3120
ggacgttggc	aaagaccagg	agttcaaggt	caaatacaag	ggtgctgggt	gtcaaggcaa	3180
agtggtatcc	aagatttgtg	gcccctcggt	tgacggcggt	ccctgcaagg	tggaagcagg	3240
ccctgggggct	gacaacagtg	tgtgtgcgtt	cctgcgccgt	gaggaaaggc	ccatagaggt	3300
ggaggtgacc	tatgacggcg	tgcccgctgc	tggcagcccc	tttctctctg	aaagcttgcc	3360
ccccaacaga	ccctagccgt	tgaaggcgtt	tggccgggg	ctgcagggag	gcagtgccgg	3420
ctcccccgcc	cgcttcacca	tcgacaccaa	gggcgccggc	acaggtggcc	tgggcctgac	3480
ggtggagggc	ccctgtgagg	cgacgtcga	ggtctggacc	aatggggatg	gcacatgttc	3540
cggtgtctac	gtgcccaacc	agccccggga	ctcaaacatt	aaatctctct	tcgctgacac	3600
ccacatccct	ggctcccat	tcaaggccca	cggtgtctcc	tgctttgacg	catccaaagt	3660
caagtgtctc	ggccccgggc	tggagcgggc	cacgcgtggg	gaggtgggac	aattccaagt	3720
ggactctctg	agcgcgggca	gcgcggagct	gacattgag	atctgctcgg	agggcgggct	3780
tcggcgagag	gtgtacatcc	aggaccaagg	tgatggcag	caacacatta	ctcacattcc	3840
ccctctgccc	ggggcctaca	ccgtcaccat	caagtacggc	ggccagcccg	tgcccaactt	3900
ccccacaaag	ctcgaggtgg	aaactgcgtg	ggacacttcc	ggtgtccagt	gtgtgtggcc	3960
tgttatgtag	gcgcagggtg	tcttcogtga	ggccaacctc	gagttcagtg	tggaagcccg	4020
ggctctgaca	cagacccggag	ggccgcacgt	caaggcccg	gtggccaacc	ccctcaggca	4080
ctgacgggag	acctacgttc	aggaccgtgg	cgatggcatg	tacaagaatg	agtagcacgc	4140
ttacgaggag	ggactgcact	cogtggacgt	ccctatgatg	ggcactcccg	tgcccagcag	4200
ccctctccag	gtgcccggtg	ccgagggtcg	cgacccctcc	cggtgtcggtg	tcacagggcc	4260
agggcatccaa	agtgccacca	ccacaagcc	caacaagtc	actgtggaga	ccaggggagc	4320
tggcaccggc	ggcctggggc	tggctgtaga	gggcccctcc	gaggccaaga	tgtctcgcat	4380
ggataaccaag	gacggcgact	gctcggtcga	gtacatccct	tataggcggt	gcacctacag	4440
cctcaacgtc	acctatggtg	gccatcaagt	gccaggcagt	cccttcagg	tcctcttgca	4500
tgaatgtaca	gatgcgtcca	aggtcaagt	gctggggc	ggcctgagcc	caggcatggt	4560
tcgtgccaac	ctctccagt	ccctccaggt	ggacacaagc	aaggctggtg	tgccccaatt	4620
cgaggtcaaa	gtgcaagggc	ccaaaggcct	ggtggagcca	gtggagctga	tagacacaacg	4680
tgatggcaac	cagacgcctca	atatgtgcct	cagcccgagaa	ggccctcaca	gcatctcag	4740
actgtatgga	gatgaagagg	taccocggag	cccttcaag	gtcaagggtc	tgccctactca	4800

tgatgccagc	aaggtgaagg	ccagtggccc	cgggtccaac	accactggcg	tgcctgccag	4860
ccctgccgtg	gagttcacca	tcgatgcaaa	ggacgcggcg	gagggcctgc	tggctgtcca	4920
gatcacaggt	cccgaaaggc	agccgaagaa	gacacacatc	caagacaacc	atgacggcac	4980
gtatcacagt	gcctacgtgc	cagacgtgac	aggtcgctac	accatcctca	tcaagtcacg	5040
tggtgacagc	atccccctct	ccccgtaccg	cgtgcgtgcc	gtgcccacgc	gggagccgac	5100
caagtgcact	gcacacagt	caatcggagg	tcacgggtca	ggtgctggca	ctgcgccacc	5160
cattcagatt	ggggaggaga	cggtgatcac	tgtggacact	aaggcgccag	gcaaaggcaa	5220
agtgacgtgc	acgtgtgca	cgctgatgg	ctcagaggtg	gatgtggacg	tgggtggagaa	5280
tgaggagcgc	actttcgaca	tcttctacac	ggccccccag	ccgggcgaat	acgtctactg	5340
tgtgcctttt	ggttgcgagc	acgtgcccac	cagccctctc	caagtgcagg	ctctggctgg	5400
ggaccagccc	tcggtgcagc	ccccctctac	gtctcagcag	ctggcccacc	agtcacacct	5460
cgccccgggc	ggccagcaga	cttgggcccc	ggagagcccc	ctggtgggtg	tcaatgggtg	5520
ggatgtgacc	agcctgaggg	cccttgacct	tgtcatcccc	tcaaccatca	agaaggggga	5580
gatcacaggc	gaggttcggg	tgccttcagg	caaggtggcg	cagccccacc	tcactgacaa	5640
caaaagacgc	accgtgaccg	tgcggtatgc	acccagcgag	gctggcctgc	acgagatgga	5700
catccgctat	gacaaatgc	acatccccag	aagcccttgc	cagttctatg	tggattacgt	5760
caactgtggc	gatctcactg	ccctatgggc	tggctcacc	catggagtag	tgacaacagg	5820
tgcacccctc	acogtcaaca	ccaaggtatc	aggagagggg	ggcctgtctc	tggccattga	5880
ggggccctcc	aaagcagaaa	tcagctgcac	tgacaaccag	gatgggacat	gcacgctgtc	5940
ctactcgctc	gtgtcgccgg	gggactacag	catttctagt	aagtacaagt	aaacagacgt	6000
cccaaggcgc	cccttcacgc	ctcgggtcac	aggtgacgac	tccatcgcta	tgtcccacct	6060
aaaggtcgcc	tctgtctcgc	acatccccat	caacatctca	gagacggatc	tcaagctcgt	6120
gaacggctct	tggttcccgc	ccctgggccc	ggagagcccc	tgtttgccta	acgcgctggc	6180
taatggccac	gtggggattt	cattcgtgcc	caaggagacg	ggggagcacc	tgggtgcatt	6240
gaagaaaaat	ggccagcacg	tggccagcag	ccccctccc	gtggtgatca	gccagtcgga	6300
aattggggat	gccagtcgtg	tccgggtctc	tggtcagggc	cttcacagaa	gccacacctt	6360
tgagcctgca	gaggttatca	tgtalaccgc	cgatgcaggc	tatgggtggc	tcaagctctc	6420
caltgaggyc	ccacagacag	tggacatcaa	cacagagggc	ctggaggacc	ggacgtgcag	6480
ggatcacatc	tgccccacag	agccaggcaa	ctacatcatc	aaatcaagt	tggccagaca	6540
gcacgtgcct	ggcagcccc	tctctgtgaa	ggtgacagcg	gagggccggg	tgaagagag	6600
catcaaccgc	agcgctcggy	ctccttcagt	ggccaacggt	ggtagtcaat	gtgacctcag	6660
ctgaaaaatc	ccgtgaaatt	gcacacagga	tatgacagcg	caggtgacca	gccatccggg	6720
caagaccatg	gagggccgaga	tctgtgaagg	ggagaaccac	acctactgca	tccgctttgt	6780
tgccgctgag	atgggcccac	acacagtcag	cgtcaagtac	aagggcccag	acgtgcgttg	6840
gagccctctc	cagttcaccc	tggggccctc	aggggaaagg	ggagcccaca	aggtcccgag	6900
tggggccctc	ggcctggaga	gagctgaagc	tggagtgcga	gccgaattca	gtatctggac	6960
ccggggaagc	ggtgctggag	gocctggccat	tgtctgcgag	ggccccagca	agggctagat	7020
ctcttttgag	gacccgaagg	acggctccgt	tgggtggctt	tatgtggtcc	aggagccagg	7080
tgactacgaa	gtctcagtca	agttcaacga	ggaacacatt	ccccagacgc	ccctcgttgt	7140
gcctgtggct	tctccgtctg	gcgacgcccc	cgccctcaet	gtttctagcc	ttcaggagtc	7200
agggctaaag	gtcaaccacg	cagccctctt	tgcagtcagc	ctgaacgggg	ccaagggggc	7260
gatcgtatgc	aaaggtgcaca	gccccctcag	agccctggag	gagtgctatg	tcacagaaat	7320
tgaccaagat	aagtatgtctg	tgcgtctcat	ccccctggat	aatggcgttt	acctgattga	7380
cgtaagttct	aacgggtaccc	acatccctgg	aagccccctc	aagatccgag	tggggagacc	7440
tgggcctgga	ggggagccag	gcttggtgtc	tgcttacgga	gcagctctgg	agccggtgtg	7500
cacaggggaac	ccagctcgagt	ctgtcgtgaa	cacagacaat	gcgggagctg	gtgcctctgt	7560
ggtgacattt	ggcgccccc	ccaaggtgaa	gatggattgc	caggagtgcc	ctgcgggtca	7620
cccgctcacc	tataccccca	tggcacctgg	cagctacctc	atctccatca	agtcagcgcg	7680
ccccctccac	atggggggca	gccccctcaa	ggccaaagtc	acagggcccc	gtctcgtcag	7740
caaccacagc	ctccacagaga	catcatcagt	gtttgtagac	tctctgacca	agggcaccct	7800
tgccccccag	catggggccc	cgggtccctg	gcctcgtcag	gccacgaagg	tgggtgccaa	7860
gggctcgggg	ctgagcaagg	cctacgttag	ccagaagagc	agcttcacag	tagactcgag	7920
caaaagcagc	aacaacatgc	tgtcgtgggg	ggttcatggc	ccaaggaccc	ctctgcgaga	7980
gatcctgggt	aagcaccgtg	gcagccggct	ctacagcgct	tctctacctg	tcaaggacaa	8040
ggggagtagc	acactggtgg	tcaaatgggg	gcacagcagc	atccccagga	gccccctacc	8100
cgttgtgggtg	ccctgagttc	ggggccctgg	ccagccggca	gcccccaagc	ctgccccgct	8160
accocaaagc	ccccggccct	tccccctcaa	ccccggccca	ggccgcccgt	gcgcgcccgc	8220
tgtcaactgca	gctgccccct	ccctgtggcg	tgcgtgcctc	cgctgcctcc	cagcagacgc	8280
gcgtacctct	cggtcttcac	ttgggcagag	ggagccattt	ggtggcgtg	cttgtctctc	8340
tgtgttctgg	gaggggtgag	ggatgggg				

<210> 74
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001456

<400> 74
 tgacctctcg gctttcactt gggcagaggg agccatttgg tggcgctgct tgtcttcttt 60

<210> 75
 <211> 1642
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001548

<400> 75
 ccagatctca gaggagcctg gctaagcaaa accctgcaga acggctgcct aatttacagc 60
 aacctagagt acaaatgggtg atgatcatca ggccaaggat agtctggagc aattgagatg 120
 tcaactttaca tgggagttat ccattgatga cgtatgaagt cctgatttag aaaacagagt 180
 cttgattcag attgaattcc tagacaccaa atacagttgt ggaatacaca acctactagc 240
 ctatgtgaaa caccatgaaag gccagaatga ggaagccctg aagagcttaa aagaagctga 300
 aaacttaagt caggaagaac atgacaacca agcaaatgtg aggagcttgg tgacctgggg 360
 caactttgcc tggatgtatt accacatggg cagactggca gaagcccaga cttacctgga 420
 caaggtggag aacatttgca agaagcttcc aaatcccttc cgctatagaa tggagtgtcc 480
 agaaatagac tgtgaggaag gatgggctt gctgaagtgt ggaggaaaag attatgaacg 540
 gggccaaggcc tgctttgaaa aggtgcttga agtggaccct gaaaaacctg aatccagcgc 600
 tgggtatgcy atctctgcct atcgcttggg tggctttaa ttagccacaa aaaatcaca 660
 gccattttct ttgcttcccc taaggcagcg tgcctcgta aatccagaca atggatatat 720
 taaggttctc cttgcccctga agcttcagga tgaaggacag gaagctgaag gagaaaagta 780
 cattgaagaa gctctagcca acatgtcctc acagacctat gtctttcgat atgcagccaa 840
 gttttaccga agaaaaggct ctgtggataa agctcttgag ttattaaaaa aggccttgca 900
 ggaacacccc acttctgtct tactgcata ccagataggg ctttgctaca aggcacaaat 960
 gatccaatc aaggaggcta caaaaggcca gcctagaggg cagaacagag aaaagctaga 1020
 caaaatgata agatcagcca tatttcattt tgaatctgca gtggaaaaaa agcccaactt 1080
 tgaagtggtc catctagacc tggccaagaat gtatataga gcaggcaatc acagaaaagg 1140
 tgaagagaat tttaaaaat tgttatgcat gaaacacagt gtagaagaaa caatgcaaga 1200
 catcattttc tactatggtc ggtttcagga atttcaaaag aaatctgacg tcaatgcaat 1260
 tatccattat taaaaagcta taaaaataga acaggcatca ttaacaaggg attaaaagta 1320
 caattctttg aagaatttgg ttttaaggaa acttcggaga aaggcattag atctggaag 1380
 cttgagcctc ctgggttctg tctataaatt ggaaggaaat atgaatgaag ccctggagta 1440
 ctatgagcgg gccctgagac tggctgtgta ctttgagaac tctgtgagac aaggtcctta 1500
 ggcaccagaa tatcagccac ttccacattt catcttaaac ttactaatc 1560
 atctttctg cttactgttt tcagaacat tataattcac tgtaatgatg taattcttga 1620
 ataataaatc tgacaaaata tt 1642

<210> 76
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001548

<400> 76
 gtatcaattc tttagaagaa ttggttttaa ggaaactctg gagaaggaga ttagatctgg 60

<210> 77
 <211> 3344
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001605

<400> 77

ggtacagctcg	cgcgctctgcg	ggaataggtg	cagcggggccc	tggcgggggg	actctgaggg	60
aggagctggg	gacggcgacc	ctaggagagt	tctttgggggt	gactttcaag	atggactcta	120
ctctaacacg	aagtgaatac	cggcagcgat	ttatagattt	cttcaagagg	aacgagcata	180
cgtagtgcca	ctcgctctgcc	accatcccat	tggatgaccc	cactttgtct	tttgccaatg	240
caggcatgaa	ccagttttaa	cccattttcc	tgaacacaa	tgacccatct	gcacccatgg	300
caagctcgag	cagagctgcc	aatacccaga	agtgcatccg	ggctgggggc	aaacaaaagt	360
acctggacga	tgtgggcaag	gatgtctatc	atcacacctt	cttcogagtg	ctgggctctt	420
ggctcttttg	agattacttt	aaggaatttg	catgtaagat	ggctctggaa	ctcctcacc	480
aagagtttgg	cattcccat	gaaagacttt	atgttactta	ctttggcggg	gatgaagcag	540
ctggctttag	agcagatctg	gaatgcaaac	agatctggca	aaatttgggg	ctggatgaca	600
ccaaaatcct	ccagagcaac	atgaaggata	acttctggga	gatgggtgac	acgggcccct	660
gtggtctctg	cagtgagatc	cactacgacc	ggattgggtg	tcgggacgcc	gcacatcttg	720
tcaaccaggga	cgaccctaat	gtgctggaga	tctggaaacct	tgtgttctat	cagtaataca	780
gggaagctgc	tggacttctg	aaacctcttc	ccaagaaaag	cattgacaca	gggatggccc	840
tgggaacgact	ggtagtctgt	ctgcagataa	agatgtccaa	ctatgacact	gacctttttg	900
tcctctactt	tgaagccatt	cagaagggca	caggtgcccg	accatacact	gggaaagtgt	960
gtgctgagga	tgcgatgggg	attgacatgg	cctacccggg	gctggctgac	catgtctogga	1020
ccatcaactg	ggcactggct	gatgggtggc	ggctgtgaca	cacagggcgt	ggatatgtgt	1080
tgagacggat	tctccgcgca	gctgtccgat	acgcccata	aaagctcaat	gccagcaggg	1140
gcttctttgc	tacggttagt	gatgtttctg	tccagtcctt	gggagatgca	tttctcagc	1200
tgaagaagga	ccagagacat	gtgaaggaca	tcattaatga	agaagaggtg	cagttttctca	1260
agactctcag	cagaaggcgt	cgactcctgg	acaggaataa	tcagagccctg	ggagacagca	1320
agaccattcc	cggaagacat	gcttggctcc	ctatgggttt	ccatgggtatc	ctgttgatct	1380
tgactggact	gattgctgaa	gagaaggggc	tggttgtaga	catggatggc	tttgaagagg	1440
agaggaaact	ggcccagctg	aaatcacagg	gcaaggagag	tggtggggaa	gacctcaatta	1500
tgctggacat	ttacgctatc	gaagagctcc	gggacggggg	tctggaggtc	acagatgatt	1560
cccctaaagta	caattaccat	ttggactcca	gtggtagcta	tgtatttgag	acacagtggt	1620
ctacggtgat	ggctctgcgc	agggagaaga	tggtctgtga	agaggtgtcc	acagggcagg	1680
agtggtgagt	gtgtctggac	aagacctgtt	tctatgctga	gcaaggaggc	cagatctatg	1740
acgaagctcta	cctgtgtgaa	gtggatgaca	cgagtgaaga	taaaaacagag	tttaccagtga	1800
agaatgtctca	gtgtccagga	gggtatgtgc	tacacattgg	aacctctgac	gttgacactga	1860
aaagtggggga	tcaggtctgg	ctgtttattg	atagaccccg	acgaagaccc	atcatgagca	1920
accacacagc	ctgacacatt	ctgaacttcg	ccctgcgctc	agtgcttggg	gaagctgacc	1980
agaagggctcc	atgtgttgct	ctgacccgcc	tcagatgttg	ctttactgct	aaggagacca	2040
tgctaccacca	acagatcaag	aaggctgaag	agattgtctaa	tgagatgatt	gagggacgcca	2100
agggcgtctca	taccacagat	tgccccctgg	cagcagcgaa	agccatccag	ggcctacggg	2160
ctgtgttttga	tgaacctat	cctgaacctg	tgcgagtctg	ctccattggg	gtcccggtgt	2220
ccgagttgct	ggtagacccc	tctggggcct	ctggctccct	gactctctgt	gagttctgtg	2280
ggggaacgca	ctctcggaac	tcagatcatg	caggagcttt	tgtgatctgt	acggaagaa	2340
ccgttgccaa	gggtatccgg	aggatgtgtg	ctgtcacagg	gtccgagggc	caggaaggccc	2400
tcaggaaagc	agagagcttg	aagaaatgtc	tctctgtcat	gggaagccaa	gtgaaggctc	2460
agactgctcc	aaacaaggat	gtgcagaggg	agatcgctga	ccttggaagg	gcctgggaca	2520
ctgcagctcat	cccccagtg	cagaaggatg	aattgcggga	gactctcaaa	tccttaagaa	2580
aggtcatgga	tgactctggac	cgagccagca	aagccgatgt	ccagaaagca	gtgttagaga	2640
agacgaagca	gttcatcgac	agcaacccca	accagcctct	tgtcatcctg	gagatggaga	2700
gcggcgctctc	agccaaggccc	ctgaatgaag	ccttgaagct	cttcaagatg	cactccocctc	2760
agactctcttc	catctctctc	acggtggaca	atgaggctgg	caagatcaog	tgccgtgtgtc	2820
aaagtccccca	gaatgcagcc	aatccggggt	taaaagccag	cgagtgggtg	cagcaggtgt	2880
caggctctgat	ggcggtgaaa	ggtggtggca	aggaatgttc	tgacacaggg	acagggaaca	2940
acgttggctg	cctgcaggag	gcgctgcagc	tgcccaattc	cttcgccag	ctgcgcctcg	3000
gggatgtaaa	gaactgagtg	gggaaggagg	aggctccac	tggaatccat	ctgcacagca	3060

```

agagctcttc atctgctaca agaacatttg aatcttggga cctttaaaga gccccctcta 3120
accacagcagt aactggaaca caettgggag cagtcctatg tctcagtgcc ccttaaattt 3180
ctgcccctgag cctctccagt cagtgccatc ggtctagaac cactaacccc gcattgctgt 3240
tgatcgtaac gctcgcatct atagataacg gctctccaga cctgagcttt ccgctgcagc 3300
aagtaggaat cgtttttgct gcagagaata aaaggaccac gtgc 3344

```

<210> 78

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001605

<400> 78

```

gccaaagagct ctctatctgc tacaagaaca tttgaatctt gggaccttta aagagccctt 60

```

<210> 79

<211> 417

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001645

<400> 79

```

acctcccaac caagccctcc agcaaggatt caggagtgcc cctcgggcct cgccatgagg 60
ctcttctgtc cgctcccggt cctgggtggtg gttctgtcga tcttcttggg aggccagcc 120
ccagccccagg ggaccccaga cgtctccagt gccttggata agctgaagga gtttggaac 180
acactggagg acaaggctcg ggaactcatc agccgcatca aacagagtga actttctgcc 240
aagatcgagg agtggttttc agagacattt cagaaaatga aggagaaact caagattgac 300
tcatgaggac ctgaagggtg acatccagga ggggcctctg aaatttccca caccagcg 360
cctgtgctga ggactccgc catgtggccc caggtgccac caataaaaaa cctaccg 417

```

<210> 80

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001645

<400> 80

```

aaacagagtg aactttctgc caagatgcgg gagtggtttt cagagacatt tcagaaagtg 60

```

<210> 81

<211> 1389

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001809

<400> 81

```

cgcggaattc tgccaagcac cggctcatgt gaggctcgcg gcacagcgtt ctctgggctc 60
cccagaagcc agcctttcgc tcccgaacc ggcagcccgga gcaggagccg tgggaccggg 120
cgccagcacc ctctcgccgg tgctatgggc ccgcgcgcgc ggagccgaaa gcccgaggcc 180
ccgaggaggc gcagcccgag ccgcgacccc acccccggcc cctcccgccg gggccctccc 240
ttaggcgctt cctcccatac acacagtcgg cggagacaag gttggctaaa ggagatccga 300
aagcttcaga agagcacaca cctcttgata aggaagctgc ccttcagccg cctggcaaga 360

```

gaaatattgt	ttaaattcac	togtgggtgt	gacttcaatt	ggcaagccca	ggccctattg	420
gccctacaag	aggcagcaga	agcattttcta	gttcatctct	ttaggagcgc	ctatctctctc	480
accttacaatg	caggcccgagt	tactctctctc	ccaaaggatg	tgcaactggc	ccggaggatc	540
cggggcccttg	aggaggagact	cggctgagct	cctgcaccca	gtgtttctgt	cagctcttcc	600
tgctcagcca	ggggggatga	taccggggac	tctccagagc	catgactaga	tccaatggat	660
tctgcgagtc	gtctgggact	ttgctgtctc	tgaacagtat	gtgtgtgttg	ctttaaatat	720
ttttcttttt	tttgagaagg	agaagactgc	atgactttcc	tctgtaacag	aggtaatatata	780
tgagacaatc	aacacccgttc	caaaggcctg	aaaataattt	tcagataaag	agactccaag	840
gttgacttta	gtttgtgagt	tactcatgtg	actatttgag	gattttgaaa	acatcagatt	900
tgctgtgga	tgggagaaaa	ggttatgtac	ttattatttt	agctctttct	gtaatatatta	960
cattttttac	catatgtaca	tttgtacttt	tattttacac	ataaggaaaa	aaataagacc	1020
acttttagca	gttgctggga	aggctgggca	tttccatcat	atagacctct	ggccttcaga	1080
gtagcctcac	cattagtggc	agcatcatgt	aactgagtgg	actgtgctgt	tcaacggatg	1140
tgtagctttt	cagaaactta	attggggatg	aatagaaaaa	ctgtgaacct	tgatgttctg	1200
gttatctcta	gtaaattcct	gtcaaaatca	attcagaaat	tctaacttgg	caactttaac	1260
attttactct	tgtaaatcat	agaagatgta	tcataacagt	tcagaatttt	aaagtacatt	1320
tcogatgctt	ttaatgggtat	ttttgtagtt	tctttgtaga	gagataataa	aaatcaaaat	1380
atttaatga	1389					

<210> 82

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001809

<400> 82

ggggatgaat agaaaacctg taagctttga tgttctggtt acttctagta aattcctgtc 60

<210> 83

<211> 2205

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001909

<400> 83

gcgcacgcgc	gccgcgccca	ogtgaccggt	cggggtgcaa	acacgcgggt	cagctgatcc	60
ggcccaactg	cggcgctcac	cggctataaa	gcgcacggcc	tccggcgacc	tctccgaccc	120
ggccgcgcgc	gccatgcagc	cctccagcct	tctgcgcctc	gcctctgccc	tgtctggctgc	180
accgcctctc	gcgctcgtca	ggatcccgct	gcacaaagtc	acgtccatcc	gccgggacat	240
gtcgcagggt	gggggctctg	tggaggacct	gattgccaaa	ggccccgtct	caaaagtactc	300
ccaggcggtg	ccagccgtga	ccgaggggcc	cattcccagag	gtgctcaaga	actacatgga	360
cgccacgtac	tacggggaga	ttggcatcgg	gcgcgcccc	cagtgcttca	cagtcgtctt	420
cgacacgggc	tcttccaacc	tgtgggtccc	ctccatccac	tgcaaacctg	tggacatcgc	480
ttgctgggat	caccacaagt	acaacagcga	caagtcacag	acctacgtga	agaatggtac	540
ctcgtttgac	atccactatg	gctcgggcag	ctctcccggt	tacctgagcc	aggacactgt	600
gtcgtgcccc	gcgcagtcag	cgctgcagc	ctctgcctgc	ggcggtgtca	aagtggagag	660
gcaggtcttt	ggggagggcca	ccaagcagcc	aggcatcacc	tcatcgcag	ccaagttcga	720
tggcatcctg	ggcatggcct	accocccgat	ctccgtcaac	aacgtgctgc	cagctcttga	780
caacctgatg	cagcagaagc	tgggtggacca	gaacatcttc	tcttctcacc	tgagcagggga	840
ccagatgcgc	cagcctgggg	gtgagctgat	gctgggtggc	acagactcca	agattacaaa	900
gggttctctg	tctcacttga	atgtcacccg	caaggcctac	tggcaggctc	acctggacca	960
ggtggagggt	gccagcgggc	tgaccctgtg	caaggagggc	tgtgaggcca	tgtgtggacac	1020
aggcacttcc	ctcatgttgg	gcccgttgga	tgaggtgcgc	gagctgcaga	aggccactcg	1080
ggccgtgcgc	ctgatcaagg	gcgagtaaat	gatccccctg	gagaagggtg	ccacctctgc	1140
cgcgatacca	ctgaactggg	gaggcaaaag	ctcaagactg	tccccagag	actacaacgt	1200
caagggtgct	caggccggga	agacccctct	cctgagcggc	tctatgggca	tggacatccc	1260

```

gccaccacgc gggccactct ggcactctgg cgacgtcttc atcgcccgct actacactgt 1320
gtttgacgct gacaacaaca ggggtgggct cgcgagggct gcccgccctt agttcccaag 1380
cggtccgcgc gcgcagcacag aaacagagga gagtcccaga gcaggaggcc cctggcccag 1440
cgggccctcc cacacacacc cacacactcg ccgcgccact gtccctggcg ccttggaagc 1500
cgggcgccca agcccgaact gctgttttgt tctgtggttt tccctccctt gggttcagaa 1560
atgctgctgt cctgtctgtc tctccatctg ttgtgtgggg gtagagctga tccagagcac 1620
agatctgttt cgtgcatctg aagacccac ccaagcttgg cagccgagct cgtgtatcct 1680
ggggctccct tcactccag ggagtcacct ccccgccctt accagcgccc cgtgggctga 1740
gccctacccc cacaccaggg cgtctcccg gccctccctt tggaacctgt cctgtcctga 1800
ggggccctct gcccaagctg gcccagctg ggctctgcca cctacctgt tcagtgtccc 1860
ggggccgttg aggatgagc cgctagaggg ctgaggatga gctggaagga gtgagagggg 1920
acaaaaccca ccttgttga gctgcaggg tgggtcgtgg actgagccag tccagggggc 1980
atgtattggc ctggaggctg ggttgggatt gggggctggg gccagccctc cctctgagct 2040
gacctctgtt gctctccctt tgggcggctg agagcccag ctgacatgga aatacagttg 2100
ttggcctccg gctccctc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 2160

```

<210> 84

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001909

<400> 84

```

tctgtttgtg ggggtagag ctgatccaga gcacagatct gtttctgca ttggaagacc 60

```

<210> 85

<211> 817

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002038

<400> 85

```

gaacggttta ctcgctgctg tgcccatcta tcagcaggct cggggctgaa gattgcttct 60
cttctctcct ccaaggtcta gtgacggagc ccgcgcgcgg cgccaccatg cggcagaagg 120
cggtatcgct ttctctgtgc tacctgctgc tcttcaacttg cagtggggtg gaggcaggta 180
agaaaagatg ctggagagcg tcggacagcg gtcocgggtt ctggaaggcc ctgaccttca 240
tgccgctcgg aggaggactc gcagtcgccc ggtctgccgc gctgggcttc accggcgccg 300
gcatcgcgcc caactcggtg gctgctcgc tgatgagctg gctcgcgctc ctgaatgggg 360
cgccgctgcc cgccgggggg ctatgtggcca cgtctgagag cctcggggct ggtggcagca 420
cggtcgtcat agtgaatatt ggtgccctga tgggctacgc caaccacaag tatctcgata 480
gtgaggagga tgaggagtag ccagcagctc ccagaacctc ttcttctctc ttggcctaac 540
ttctccagtt aggatctaga accttgccct tttttttttt tttttttttt tttagagatg 600
gttctcacta tattgtccag gctagagctg agtggtctatt cacagatgag aacatagtac 660
actgcagcct ccaactccta gctcgaagtg atctcctgt ctcaacctcc caagtaggat 720
tacaagcatg cgccagcatg gccagaaatc cagaacttgg tctatcactc tccccacaa 780
ctagatgtg aaacagaat aaacttcacc cagaaaaa 817

```

<210> 86

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002038

<400> 86
agctcccaga acctcttctt cctctcttgc ctaactcttc cagtttagat ctagaacttt 60

<210> 87
<211> 1283
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002046

<400> 87
ctctctgctc ctctctgtcg acagtcagcc goactcttct ttgcgtgcgc agccgagcca 60
catcgctcag acaccatggg gtcggagctc acggatttgg tegtatttgg 120
cgccgtggtca ccagggtgtc ttttaactct ggttaagtgg atattgttgc catcaatgac 180
cccttcattg acctcaacta catgggtttac atgttccaat atgattccac ccattggcaaa 240
ttccattgga ccgtcaaggc tgagaacggg aagcttgtca tcaatggaaa tcccatcacc 300
atcttccagg agcgagatcc ctccaaaatc aagtggggcg atgctggcgc ttagtagctc 360
gtggagtcca ctggcgctct caccaccatg gagaaggctg gggctcattt gcagggggga 420
gccaaaaggg tcatcatctc tgccccctct gctgatgcc ccattgttct catgggtgtg 480
aaccatgaga agtatgacaa cagcctcaag atcatcagc atgcctcctg caccaccaac 540
tgcttagcac ccttgggcaa ggtcatccat gacaacttgg gtatcgtgga aggactcatg 600
accacagtcc atgccatcac tgccaccag aagactgtgg atggccctc cgggaaactg 660
tggcgtagtg gcccgggggt tctccagaac atatccctg cctctactgg cgctgccaaag 720
ctgtgtggga aggtcatccc tgagctgaac ggggaagctc atggcatggc ctctccgttc 780
ccactgccca acgtgtcagt ggtggacctg acctgccgtc tagaaaaacc tgccaaatat 840
gatgacatca agaaggtggt gaagcaggcg tcggagggcc cctcaaggc catcctgggc 900
tacaactgagc accaggtggt ctctctgac ttcaacagcg acaccacac ctccaccttt 960
gacgtctggg ctggcattgc cctcaacgac cactttgtca agctcattc ctggtagtac 1020
aacgaatttg gctcacagcaa cagggtgggt gacctcatgg ccacatggc ctccaaggag 1080
taagaccctt ggaccaccag cccacagcaag agcacaagag gaagagagag acctcactg 1140
ctggggagtc cctgccacac tcagtccccc accacactga atctccccct ctacacgttg 1200
catgtagac cccttgaaga gggggagggc ctaggaggac gcacctgtgc atgtaccatc 1260
aataaagtac cctgtgctca acc 1283

<210> 88
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002046

<400> 88
ctcaacgacc actttgtcaa gctcatttcc tggtagaca acgaatttgg ctacagcaac 60

<210> 89
<211> 1610
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002061

<400> 89
ggcagcaggc tgcggccgca gtagccggag ccggagccgc agccaccggt gccttctctt 60
cccgcctgcc ccacagccgc gtcgggctc cctcggggcc gagcgagac caggtctcag 120
ccgcgcggcg ccggcagcct cgcgctccct ctccgggtctc tctcgggctc cgggcaccgc 180
gtcctgttgg ggcgcgcctg cctgcccgcc cgcccgacgc ccttgcctg ccggccctctg 240
ggcggccctg gccatgggca ccgacagccg ccggggcaag cgctcctgg gcggggcccg 300
caccctgcac ctgcagacgg ggaacctgct gaactggggc cgctcgcca agaagtggcc 360

gtccacgcac	agcgaggagc	ttcatgattg	tatccaaaaa	acottgaatg	aatggagttc	420
ccaaatcaac	ccagatttgg	tcagggagtt	tccagatgtc	ttggaatgca	ctgtatctca	480
tcgtgtagaa	aagatataatc	ctgatgaaag	agaagaatgt	aaagtctctg	caaaaactgtt	540
cattcatgaa	tcaaactctt	catcatcaac	tagaagtgca	gttgacatgg	cctgtcagtt	600
ccttggagtt	gcacagctgg	attctgtgat	cattgcttca	cctcctattg	aagatggagt	660
taactctttc	ttggagcatt	tacagcctta	ctggggaggaa	ttagaaaaact	tagttcagag	720
caaaaagatt	gttgccatag	gtacctctga	tctagacaaa	acacagcttg	aacagctgtg	780
tcagtgggca	caggttaaac	caaatagtaa	ccaagttaat	cttgctctct	gctgtgtgat	840
gccaccagat	ttgactgcat	ttgctaaaca	atttgacata	cagctgttga	ctcacatga	900
tccaaaagaa	ctgctttctg	aagcaagttt	ccaagaagct	cttcaggaaa	gcattcctga	960
cattcaagcg	cacgagtggt	tgccgctgtg	gctactcggt	tattcggtca	ttgtgaaaag	1020
tagaggaaat	atcaaatcaa	aaggctacat	tttacaagct	aaaagaaggg	gttccttaact	1080
gaactaggag	cataacttac	ctgtaatctt	cttcaatgt	agagaaaatt	gagatgtgta	1140
aaactctagt	actgctgtga	aatgggtgtc	ttgaggcaga	tattctttcg	tcattattga	1200
cagtatgttg	tcgtctcaag	tttaataact	tatctgtcct	ccatatcaat	ccattctcat	1260
gaacctctgt	attgctttcc	ttaaactatt	gttttcta	tgaaattgtc	tataaagaaa	1320
tactttgcaa	tatatatttc	ctttattttt	atgactaata	taaatcaaga	aaatttgttg	1380
ttagatatat	tttggcctag	gtatcagggg	aatgtatata	catatttttt	atttccaaaa	1440
aaaattcatt	aattgcttct	taactcttat	tataaccaag	caatttaatt	acaattgtta	1500
aaactgaaat	actggaagaa	gataattttc	ctgctattga	tgagatatat	cagagttaact	1560
ggagtagctg	ggatttacta	gtagtgtaaa	taaaattcac	tcttcaatca	1610	

<210> 90

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002061

<400> 90

ctgacttagg agcataaact acctgtaatt tccctcaata tgagagaaaa ttgagatgtg 60

<210> 91

<211> 873

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002106

<400> 91

cgagctttga	atcgcggtgc	gacgaaggag	taggtgtgtg	gatctccacc	tggtgctgat	60
tagccttttc	tcgtccttgc	ttgcttgagc	ttcagcggaa	ttcgaaatgg	ctggcggtga	120
ggctggaagag	gaactccggaa	aggccaagac	aaaggcggtt	tccgcgtcgc	agagagccgg	180
cttgcaagtct	ccagtgggcc	gtattcatcg	acacctaaaa	cttaggacga	ccagctatgg	240
acgtgtgggc	ggactgccc	ctgtgtacag	cgcagccatc	ctggagtacc	tcaccgcaga	300
ggtaacttgaa	ctggcgaggaa	atgcatacaa	agacttaag	gtaaagcgtg	ttaccctctg	360
tcaacttgcaa	ctgtctattc	gtggagatga	agaattggat	tctctcatca	aggctacaat	420
tgctgtgggt	gtgtgcatct	cacacatcca	caaatctctg	attgggaaga	aaggacaaca	480
gaagactgtc	taaaggatgc	ctggatttct	tgttatctca	ggactctaaa	tactctaaaca	540
gtgtccagtg	gttggtgatt	ccagtggaact	gtatctctgt	gaaaaacaca	attttgcctt	600
tttgttaatt	tatttgagca	agttggaagt	ttaattagct	ttccaaccaa	ccaaatttct	660
gcattctgagt	cttaaccata	tttaagtgtt	actgtggctt	caagaagctt	attgtctctg	720
aagtatgtgg	ttttgattga	gttgactgtt	tttaaaaaac	tggttggatt	ttaattgtga	780
tcgagaagtt	atagtaacaa	acatttgggt	ttgtcacagc	attatttcca	ctctggtgga	840
taagtccaat	aaaggtcata	tcccaacta	aaa	873		

<210> 92

<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_002106

<400> 92
cgagtccttaa ccatatttaa gtgttactgt ggcttcaag aagctattga ttctgaagta 60

<210> 93
<211> 4204
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002205

<400> 93
caggacaggg aagagcgggc gctatgggga gccggcagcc agagtccctt ctccacggcg 60
tgacagctgcg ctggggccccc cggcgccgac ccccgctcgt gccgctcgtg ttgctgctcg 120
tgccgcgcgcc acccagggttc gggggcttca acttagacgc ggaggcccca gcagtactct 180
cggggccccc gggctccttc ttcggtattct cagtggagtt ttaccggccg ggaacagacg 240
gggtcagtggt gctgggtggga gcaccaagg ctaataccag ccagccagga gtgctgcagg 300
gtgggtcgtgt ctacctctgt ccttgggggt ccagccccac acagtgcaac ccatttgaat 360
ttgacagcaa aggcctctcg cctcctggagt cctcactgtc cagctcagag ggagaggagc 420
ctgtggaagta caagtctctg cagtgtgttc gggcaacagt tcagacccat ggctcctcca 480
tcttggcatg cgctccaactg tacagctggc gcacagagaa ggagccactg agcgaccccg 540
tgggcactctg ctacctctcc acagataaact tcacccgaat tctggagtat gcacctctgc 600
gctcagattt cagctgggca gcaggacagg gttactgcca aggaggtctc agtgccgagt 660
tcaccaagac tggcgcgtgtg gtttaggtg gaccaggaag ctatttctgg gcacctctgc 720
tctgtctcgc cactcaggag cagattgcag aatcttatta ccccgatcac ctgatcaacc 780
tggttcaggg ccagctgcag actcgccagg ccagttccat ctatgatgac agctacactag 840
gatactctgt ggctgttgtt gaattcagtg gtgatgcac agaagacttt gttgctggtg 900
tgccccaaag gaacctcact taaggctatg tcacctcctt taatggctca gacattcgat 960
ccctctcaaa cttctcaggg gaacagatgg cctcctactt tggctatgca gtggccgcca 1020
cagacgtcaa tggggacggg ctggatgact tgctgggtgg ggcaacctg ctcatggatc 1080
ggacccctga cggggcgccct caggaggtgg gcagggtcta cgtctacctg cagcaaccag 1140
ccggcataga gcccaagccc acccttacc ctaactggcca tgatgagttt ggccgatttg 1200
cgagctcctt gacccccctg ggggacctgg accagatgg ctacaatgat gtggccaatc 1260
gggtcctcctt tgggtgggag acccagcagg gagtatgttt tgaatttctt gggggccagc 1320
gaggggtggg ctctaaagct tccaggttc tgacgcccc gtgggcagcc agccaacccc 1380
cagactctct ttgctctgcc cttcgaggag ccgagacact ggaatggcaat ggatatcctg 1440
atctgatttg ggggtccttt ggtgtggaca aggcgtgtgt atacaggggc gcccactcg 1500
tgctcogtag tgctccctc accatctcc ccggcatgtt caaccagag gagcgagct 1560
gcagcttaga ggggaacctg gtggcctgca tcaaccttag cttctgcctc gatctctctg 1620
gaaaacacgt tgcgtactcc attggtttca cagtggaaat tcagctggag tggcagaagc 1680
agaaggaggg ggtacggcgg gcactgttcc tggcctccag gcaggcaacc ctgaccagca 1740
ccctgctcat ccagaatggg gctcgagagg attgcagaga gatgaagatc taactcagga 1800
acagtgacga atttcgagac aaactctcgc atgtctcaat cgctctcaac ttctccttgg 1860
accoccaaag ccagtgagac agccacggcc tcaggccagc cctacattat cagagcaaga 1920
gccgcttaga ggacaaggct cagactcttc tggaactgtg agaagacac actctgtgtc 1980
ctgacctgca gctggaagtg tttggggagc agaaccatgt gtacctgggt gacaagaatg 2040
ccctgaacct cactttccat gcccagaatg tgggtgaagg ttggccttat ggcgtgagc 2100
ttcgggtcac cgcccccca gaggctgagt actcaggact cgtcagacac ccagggaact 2160
tctccagcct gagctgtgac tactttgccg tgaaccagag ccgctctgct gtgtgtgacc 2220
tgggcaaccc catgaaggca ggagccagtc tgggggtgg ccttcggttt acagtccctc 2280
atctccggga cactaagaaa accatccagt ttgaattcca gatcctcagc gaattctcca 2340
acaactcgca aagcgacgtg gtttctcttc ggctctccgt ggaggctcag gccacggcta 2400
tgcgaacggt tgtctccaa cctgagcgag tctgaatccc agtaaacgac tggcatcccc 2460
gagaccagcg tcagaaggag gaggacctgg gacctgtctt ccacctatgc tatgatctca 2520

```

tcaaccaagg cccagctcc attagccagg gtgtgtgga actcaetgt cccaggttc 2580
tggaaggtca gcagctccta tatgtgacca gaggtaacgg actcaactgc accaccaatc 2640
accccattaa ccaaaagggc ctggagttgg atccccgagg tccccctgcac caccagcaaa 2700
aacgggaagc tccaagccgc agctctgctt cctcgggacc tcagatcctg aaatgcccg 2760
aggtcgagtg ttccaggctg cgtctgtgagc tcggggccctc gcaccaacaa gagagccaaa 2820
gtctcgagtt gcatttccga gtctggggcca agactttctt cgagcgggag caccagccat 2880
ttagctcgca gtgtgaggct gtgtacaaa cctgaagat gccctaccga atcctgcctc 2940
ggcagctgcc ccaaaaaagag cgtcaggtgg ccaagctgt gcaatggacc aaggcagaag 3000
gcagctatgg cgtccactg ttgatcatca tctagcctat cctgtttggc ctctgtctcc 3060
tagttctact ctctacatc ctctacaagc ttggattctt caaacgctcc ctcccatatg 3120
gcaccgccat ggaaaaagct cagctcaagc ctccagccac ctctgatgcc ttagtccctc 3180
caatttcaga ctccccattc tgaagaacca gtccccccac cctcattcta ctgaaaaagg 3240
ggggtctggg tactttctga aggtgctgac ggccagggag aagctcctct cccagcccca 3300
gagacatact tgaagggcca gagccagggg ggtgaggagc tgggatatcc tccccccat 3360
gcactgtgaa ggaccttgt ttacacatac cctcttcctg gatgggggaa ctcatgcca 3420
gggacagagg cccagcctcc ctgaagcctt tgcattttgg agagtttctt gaaacaactg 3480
gaaagataac taggaaatcc attcacagtt ctttgggcca gacatgccac aaggacttcc 3540
tgtccagctc caactcgcaa agatctgtcc tcagccttgc cagagatcca aaagaagccc 3600
ccagttaagaa cctggaactt ggggagttaa gacctggcag ctctggacag ccccaactgc 3660
tgtgggccaac aaagaacact aactatgcat ggtgccccag gaccagctca ggacagatgc 3720
cacaaggata gatgctggcc caggggccaga gccagctcc aaggggaatc agaatcaaa 3780
tgtgggcaga tcagcctgg ggtctggagt tgatttgaa cccagactca gacattggca 3840
cfaatccagg cagatccagg actatatttg ggctgctcc agacctgtac ctggagagccc 3900
agttcaccct gattttaggag aagccaggaa ttccccagga cctgaagggg ccatgatggc 3960
aaagcatctg gaacctcagc ctggccacag acagggccctc ctgttcccc agagaagggg 4020
gagccactag tcttgggctc gcgaatttg ggttctgct gccagctgca ctgatgctgc 4080
cctcatctc tctggccaac ccttccctca ccttggcac agacacccag gacttattta 4140
aactctgttg caagtgcatt aaatctgacc cagtggcccc actgaccaga actagaaaaa 4200
aaaa 4204

```

```

<210> 94
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002205

```

```

<400> 94
ttggcaccag acaccagga cttattttaa ctctgttgca agtgcaataa atctgaccga 60

```

```

<210> 95
<211> 1976
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002266

```

```

<400> 95
gccacaggt ctttgagctg agtcgaggtg gacctttga acgcagtcgc cctacagccg 60
ctgtatcccc ccgcatcgcc tcccgtagaa gccagggccc gctctcgagc tttctccctt 120
tgtctcataa ccatgtccac caacgagaat gctaatacac cagctgccgc tttccacaga 180
ttcaagaaca agggaaaaga cagtacagaa atgaggcgctc gcagaataga ggtcaattgt 240
gagctaggga aagctaaaga ggatgaccag atgctgaaga gaggaaatgt aagctcattt 300
cctgatgatg ctactctccc gctgcaggaa aaccgcaaca accagggcac tgtaaattgg 360
tetgttgatg acattgtcaa aggcataaat agcagcaatc tggaaaaatca gctccaagct 420
actcaagctg ccaggaaact actttccaga gaaaaacagc ccccataga caacataatc 480
ggggtctggt tgattccgaa atttgtgtcc ttctgggcca gaactgatgt tagtccattt 540
cagtttgaat ctgcttgggc actcaactaac attgcttctg ggacatcaga acaaaccaag 600
gctgtggttag atggagggtg catcccagca ttcattttctc tgttggcatc tccccatgct 660

```

```

cacatcagtg aacaagctgt ctgggctcta ggaacattg caggtgatgg ctacagtttc 720
cgagactgtt ttattaagta cgttgcaagt gaccactgt tggctctcct tgcagttcct 780
gatagtctat cttagcatg ttggtactta cgttaactta cctggacact tcttaactct 840
tgccgcaaca agaactcctgc acccccgata gatgctgttg agcagatctt tcttaccctta 900
ttctcggtctg tgcctcatga tgatccagaa gtgttagcag atacctgctg ggctattttcc 960
taacctactg atggtccaaa tgaacgaatt ggcattggtg tgaaaacagg agttgtgccc 1020
caacttgtga agcttctagg agcttctgaa ttgccaatg tgactcctgc cctaaagagcc 1080
atagggaata ttgtcaactgg tacagatgaa cagactcagg ttgtgattga tgcaggagca 1140
ctcgccgtct ttccagcct gctcaccaac cccaaaacta acattcagaa ggaagctacg 1200
tgacaatgt caaacatcac agccggccgc caggaccaga tacagcaagt tgtgaatcat 1260
ggattagtcc cattccttgt cagtgttctc tctaaggcag attttaagac acaaaaggaa 1320
gctgtgtggg ccgtgaccaa ctataccagt ggtggaacag ttgaacagat tgtgtacctt 1380
gttccactgt gcataataga accgttgatg aaacctttaa ctgaaaagaa taccagagtt 1440
attctggatta tctctggatgc catttcaaat atotttcagg ctctgagaa actaggtgaa 1500
actgagaac ttagtataat gattgaagaa tgtggaggct tagacaaaat tgaagctcta 1560
caaaaccatg aaaatgagtc tgtgtataag gcttctgtta gcttaattga gaagtatttc 1620
tctgtagagg aaagaggaaa gtaccagaaa ctacctctga aggtacacgt 1680
ttccaaagtc aggatggggc tctctgggac tttaactttt agatcatgta gctgagacat 1740
aaatttgttg tctactacgt ttggtatttt gtcttattgt ttctctacta agaacctctt 1800
cttaaatgtg gtttgttact gtgacacttt ttacactgaa actatacttg aacagttcca 1860
actgtacata ctaactgtat gaagcttgtc ctctgactag gtttctaatt tctatgtgga 1920
atttcctatc ttgcagcatc ctgtaataaa acattcaagt ccacctctaa aaaaaa 1976

```

<210> 96

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002266

<400> 96

tgagtctgtg tataaggett cgtaaagctt aattgagaag tatttctctg tagagaaga 60

<210> 97

<211> 1145

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002346

<400> 97

```

gtcccgcca gcccggtcc agagcgcgcg aggttcgggg agctccgccca ggctgctggt 60
acctgcgtcc gcccgcgag caggacaggc tgccttgggt tgtgacctcc aggcaggagc 120
gccatcctct ccagaatgaa gatcttcttg ccagtgcgtc tggctgccct tctgggtgtg 180
gagcgagcca gctcctgatg gtgcttctcc tcttgaacc agaagagcaa tctgtactgc 240
ctgaagccga ccatctgctc cgacaggagc aactactgag tgactgtgtc tgctagtgtc 300
ggcatctggg atctcgtgac atttggccac agcctgagca agacctgttc ccggcgctgc 360
ccatcccgag aaggcgtcaa tgttggtgtg gcttccatgg gcatcagctg ctgccagagc 420
ttctctgcca atttcagtgc ggccgatggc gggtgcgggg caagcgtcac cctgctgggt 480
gcccgggtct tgctgagcct gctgccggcc ctgctgcggt ttggccctct accgccaga 540
ccctgtcccc cgatccccca gctcaggaa gaaagcccg cctttctgg attccacagt 600
gtatggagac cctgtactcc tcaagtgcct gatctgtgcc ctgtgtccca ggtagagccc 660
accccctgca cctcaccctg cccagcccc tgcctctgcc caagtggggc agctgccctc 720
acttctgggg ttgatgatgt gaccttcttt ggggagctgc ggaagggagc aggtttccct 780
ggagctcttac ggtccaacat cacaccaagt ccatggaca tgcctgacag tctccaggg 840
agacctgttc agtagggatg tgtgcctggc tgtgtacgtg ggtgtcaggt gcaagtgtga 900
gcaagtgttg gctctgggg gccatgtttg gggagggagg tgtgccagca gctcggagag 960
cctcagctccc tgtagccccc tgccctggca cagctgcagt cactcaagg cagcctcttg 1020
gggggtgggg ttcttgccac tcccggtctt aggcctgtgc caaatccagc cagtctgtcc 1080

```

ccagcccacc cccacattgg agccctcctg ctgctttggt gcctcaaata aatacagatg 1140
tccc 1145

<210> 98
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002346

<400> 98
ggttccctgg agtcttaccg tccaacatca gaccaagtcc catggacatg ctgacagggt 60

<210> 99
<211> 1390
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002358

<400> 99
gggaagtgtct gttggagcgg ctgtggttgc tgtccgcgga gtggaagcgc gtgcttttgt 60
ttgtgtccct gcccatggcg ctgcagctct cccgggagca gggaaacacc ctgcgcggga 120
gcgcggaat ctgtgcgcgag ttcttctcat tcggcatcaa cagcatttta tatcgcgtg 180
gcataatcc atctgaaacc ttactcggag tgcagaaata cggactcaoc ttgcttgtaa 240
ctactgatct tgagctcata aaatacctaa ataattgtgt ggaacaactg aaagattggt 300
tatcaaatgt ttcagttcag aaactggtg tagttatctc aaatattgaa agtggtaggg 360
tcctggaag atggcagttt gatattgagt gtgacaagac tgcaaaagat gacagtgcac 420
ccagagaaaa gtctcagaaa gctatccagg atgaaatccg ttcaagtatc agacagatca 480
cagctacggg gacatttctg ccactgttgg aagtttcttg ttcatattgat ctgctgattt 540
atacagacaa agatttgggt gtacctgaaa aatgggaaga gtcgggacca cagtttatta 600
ccaattctga ggaagtccgc ctctgttcat ttactactac aatccacaaa gtaaatagca 660
tggtggccta caaaattcct gtcaatgact gaggatgaca tgaggaaaa aatgtaattg 720
taattttgaa atgtggtttt cctgaaatca ggtcatctat agttgatatg ttttatttca 780
ttggttaatt ttacatgga gaaaaccaa atgatactta ctgaactgtg tgtaattgtt 840
cctttatttt ttgtgtacct atttgactta ccatggagtt aacatcatga atttattgca 900
cattgttcaa aaggaaaccg gaggtttttt tgtcaacatt gtgagtata ttcccttgaa 960
gatagtaact gtatgtgaa aaacttgtgc tataaageta gatgctttcc taaatcagat 1020
gttttggtca agtagtttga ctacgtatag gtaggagatg atttaagtat aaaaacaac 1080
aaaggaagtc taaattatca gaatctttgt taaggtctctg aaagtaactc ataactata 1140
aacaatgaaa tattgtgtga tagctccttt tgaccttcat ttcatgtata gttttcccta 1200
ttgaatcagt ttccaattat ttgactttaa ttatgttaac ttgaacctat gaagcaatg 1260
ataattgtac tgttttaagt tctgtgtatc agaactctta aaaaagtgtt ttcattgtgt 1320
ttataaaatc aagtttttaag tgaaagttag gaaataaagt taagtttgtt ttataaaaaa 1380
aaaaaaaaa 1390

<210> 100
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002358

<400> 100
atgctttcct aaatcagatg ttttggtcaa gtagtttgac tcagtatagg tagggagata 60

<210> 101
<211> 1821

<212> DNA
<213> Homo sapiens

<300>
<308> NM_002422

<400> 101
acaagaggc aggcgaagaca gcaaggccta gagacaacat agagctaagt aaagccagtg 60
gaaatgaaga gtcttccaat cctactgttg ctgtgcgtgg cagtttgctc agcctatcca 120
ttgtagggag gtgcacgggg tgaggacacc agcatgaacc ttgttcagaa atatctagaa 180
aactactacg acctcaaaaa agatgtgaaa cagtttgta ggagaaagga cagtggtcct 240
gtgttataaa aaatccgaga aatgcagaag ttccctggat tggaggtgac ggggaagctg 300
gactccgaca ctctggaggt gatgcgcaag ccaggtgtg gagttcctga tgttggtcac 360
ttcagaacct ttccctggcat ccggaagtgg aggaaaaccc acottacata caggattgtg 420
aattatacac cagatttgcc aaagatgct gttgattctg ctgttgagaa agctctgaaa 480
gtctgggaag aggtgactcc actcacattc tccaggtgt atgaaggaga ggtcgatata 540
atgatctctt ttgcagttag agaactgga gacttttacc cttttgatgg acctggaaa 600
gttttggccc atgcctatgc cctctggcca gggattaatg gagatgcccc ctttgatgat 660
gatgaacaat ggacaaggga tacaacaggg accaatttat ttctcgttgc tgcctatgaa 720
attggccaact cctcgtgtct ctttactca gccaacactg aagctttgat gtaccactc 780
atacactcac tcacagacct gactcgttcc cgctgtctc aagatgatat aaatggcatt 840
cagtccctct atggacctcc cctgactcc cctgagacc cctcgtgtacc caccggaact 900
gtccctccag aacctgggac gccagccaac tgtgatcctg ctttgtcctt tgaatgctgc 960
agcactctga ggggagaat cctgatcttt aaagacaggc acttttggcg caaatccctc 1020
agggaacttg aacctgaatt gcatttgatc tcttcatttt ggccatctct tcttcaggc 1080
gttgtagcgc catatgaagt tactagcaag gacctgttt tcaattttaa aggaatacaa 1140
ttctgggcca tcagagaaa tgaggtacga gctggatacc caagaggcat ccacacccta 1200
ggtttccctc caaccgtgag gaaatcgtat gcagccattt ctgataagga aaagaacaaa 1260
acataatttt ttgtagagga caaatactgg agatttgatg agaagagaaa ttocatggag 1320
ccaggctttc ccaagcaaat agctgaagac ttccaggga ttgactcaaa gattgatgct 1380
gtttttgaag aattttgggt cttttatttc ttactggat cttcacagtt ggaatttgac 1440
ccaaatgcaa agaagtgcac acacactttg aagagtaaca ctggccttaa ttgtgaaag 1500
agatatgtag aaggcacaat atgggcactt taatgaagc taataattct tcacctaaat 1560
ctctgtgaat tgaatgttc gttttctcct gcctgtgctg tgactcgagt cacactcaag 1620
ggaaactgag cgtgaatctg tatcttgccg gtccattttta tgttattaca gggcatcaa 1680
atgggtcgtc gcttagcttg cacctgtgca catagagtga tctttcccaa gagaaggga 1740
agcactcgtg tccaacagac aagtactgtt atctgtgtag actatttgct tatttaataa 1800
agacgttttg tcagttgtt t 1821

<210> 102
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002422

<400> 102
tgtagaaggc acaatatggg cacttttaaat gaagctaata attcttcacc taagtctctg 60

<210> 103
<211> 2787
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002462

<400> 103
agagcggagg cgcactcca gcaactgcga gggaccgctt tggaccgag ttgcgggcca 60
ggaaatccag tgcacgggtg gacacgcctc cctcgcgccc ttgcgcacca cctgctcacc 120

caagctcaggg	gctttggaat	tctgtggcca	cactgcgagg	agatcggttc	tgggtcggag	180
gctacaggaa	gactcccaact	ccctgaaatc	tggagtgaag	aacgcgcgcca	tcaccagacc	240
attccaaggga	gggtgcaggag	aacagctctg	tgataccatt	taacttggttg	acattcattct	300
tattttgaagg	aacgttatatt	agagcttact	ttgcacaaaga	ggaagatggt	tggttccgaa	360
gtggacatcg	caaaagctga	tccagctgct	gcataccacc	ctctattact	gaatggagat	420
gctactgtgg	ccgcagaaaa	tccaggctcg	gtggctgaga	acaacctgtg	gagccagtat	480
gaggagaagg	tgccgccccg	catcgacctc	attgactccc	tgccgggtct	agggtggag	540
caggacctgg	ccctgcacgc	catcgccgtc	atcggggacc	agagctcggg	caagagctcc	600
gtgttggagg	caactgtcagg	agttgccctt	cccagaggca	gcgggatcgt	gaccagatgc	660
ccgctgtgtc	tgaacttgaa	gaaactttgt	aacgaagata	agtggagagg	caaggtcaagt	720
taccaggact	acagatttga	gatttcggat	gcttcagagg	tagaaaaaga	aattaataaa	780
atcagctccc	gagatgtccc	ggatctgact	ctaatagacc	ttcctggcat	aaccagagtgc	840
gctgtgggca	atgacgtctc	tgacatttgg	tataagatca	agacactcat	caagaagtac	900
atccagaggc	aggagacaat	cagcctgggt	gtggtcccca	gtaattgtga	catcgccacc	960
acagaggctc	tcagcatggc	ccaggagggt	gaccocaggg	gagacaggac	catcggaatc	1020
ttgacgaagc	ctgatctggt	ggacaaagga	actgaagaca	aggttgtgga	cgtggtgcgg	1080
aacctcgtgt	tcacactgaa	gaagggttac	atgattgtca	agtgcocggg	ccagcaggag	1140
atccaggacc	agctgagcct	gtccgaagcc	ctgcagagag	agaagatctt	ctttgagaac	1200
caccctattt	tcagggatct	gctggaggaa	ggaaaggcca	cggttccctg	cttgcgcagaa	1260
aaactccga	ccagcgtcat	cacacatac	tgtaaatctc	tgccctgtgt	agaaaaatcaa	1320
atcaaggaga	ctccaccagag	aataacagag	gagctacaaa	agtattggtg	cgacataaccg	1380
gaagacgaaa	atgaaaaaat	gttcttctcg	atagataaaa	ttatgcctt	taatcaggac	1440
atctactgtc	tcatgcgaag	agaggaaact	gtaggggagg	aagacattcg	gctgtttacc	1500
agactccgac	acgagttcca	caaattggagt	acaataattg	aaaacaaatt	tcgaagaaggc	1560
caataaaatt	tgagttagaaa	aatccagaaa	tttgaataat	agtatcgtgg	tagagagctgc	1620
ccaggctttg	tgaattacag	gacattttag	acaactcgtga	aacagcaaat	caagggactg	1680
gaagagccgg	ctgtggatgat	gctacacacc	gtgacggata	tggtcgcgct	tgctttcaca	1740
gatgtttcga	taaaaaattt	tgaagagttt	tttaacctcc	acagaaccgc	caagtccaaa	1800
attgaagaca	ttagagcaga	acaagagaga	gaaggtgaga	agctgatccg	ccctccattc	1860
cagatggaac	agatgtgcta	ctgccaggac	caggtatata	ggggtgcat	gcagaaaggtc	1920
agagagaagg	agctgggaaga	agaaaaaga	aagaaatcct	gggattttgg	ggctttccag	1980
tcacagctcg	caacagactc	ttccattggag	gagatctttc	agcaacctgat	ggcctatcac	2040
caggaggcca	gcaagcgcac	ctccagccac	atccctttga	tcactccagt	cttcactgtc	2100
cagacgtacg	gccagcagct	tcagaaggcc	atgctgcagc	tcctgcaggga	caaggacacc	2160
tacagctggc	tcttgaagga	cgggagcgac	accagcgaca	agcggaagt	cctgaaggag	2220
cggcttgacg	ggctgacgca	ggctcggcgc	cggcttgccc	agttccccgg	ttaccacac	2280
ctgttccagc	cccgtagacg	tgcacgcaca	ctgtctgccc	ccgttccggg	cagtcactgt	2340
gactgcagac	ttagttgctc	agtagtcaga	ctggatagtc	cgtctctgct	tatccgttag	2400
ccgctgtgat	ttagcaggaa	cgtgtgagag	cagttttggt	tctagcatga	agacagagcc	2460
ccaccctcag	atgacatga	gctggcggga	ttgaaggatg	ctgtctctgt	actgggaaag	2520
ggattttcag	ccctcagaat	cgtccacact	tgacgctctc	cccttctctg	tattctcaga	2580
aactgcacaca	tgtctgaacat	cacagcttat	ttctctattt	ttataatgtc	ccttcacaaa	2640
cccagtgttt	taggagcatg	agtgcgctgt	gtgtgcgtcc	tgctcgagcc	ctgtctcctc	2700
tctctgtaat	aaactcattt	ctagcag	2787			2760

<210> 104

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002462

<400> 104

actgcacat gctgaacatc acagcttatt tctctatttt tataatgtcc cttcacaac 60

<210> 105

<211> 2808

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002759

<400> 105

```

cgggcgccggt  cgggcgccagt  tgctcatact  ttgtgactgt  cggtcacagt  ggcattcagc  60
tccacactgtg  gtgagaaccac  aggcacgaca  agcatagaaa  catcctaaac  aatcttcac  120
gaggcatcgga  ggtccatccc  aataaaaatc  aggagaccc  ggctatcata  gaccttagtc  180
tcgctcggtta  tactcgctgt  ctgtcaacca  gcgggttgact  ttttttaagc  ctcttttttt  240
ctcttttacc  agtttttggg  gcaaatccag  ttgctctcc  tggatttgta  ttcgttaagt  300
acctcaaac  tttagcagtt  ctccactctg  actcaggttt  gcttctctgg  cggttctcag  360
aatcaacatc  cacacttccg  tgattatctg  cgtgcatttt  ggacaaagct  tccaaaccag  420
atacgggaag  aagaaatggc  tgggtgactt  tcagcagggt  tcttcattga  ggaacttaat  480
acataccgct  agaagcgagg  agtagtactt  aaatatcaag  aactgcctaa  ttcaggacct  540
ccacatgata  ggaggtttac  atttcaagtt  ataatagatg  gaaggaatt  tccagaaggt  600
gaaggtagat  caaagaagga  agcaaaaaat  gccgcagcca  aattagctgt  tgagataact  660
aataaggaaa  agaaggcagt  tagtcttcta  ttattgacaa  caacgaatc  ttcagaagga  720
ttatccatgg  ggaatttcat  aggccttatt  aatagaattg  cccagaagaa  aagactaact  780
gtaaattatg  aacagtggtg  atcgggggtg  catgggccag  aaggatttca  ttataaatgc  840
aaaatgggac  aagaaagaata  tagtattggt  acaggttcta  ctataacagga  agcaaaaaca  900
ttggccgcta  aacttcgata  tcttcagata  ttatcagaag  aaacctcagt  gaaactcgac  960
tacctgtcct  ctggttcttt  tgctactacg  tgtgactccc  aaagcaactc  tttagtgacc  1020
agcactctcg  ctctgtaatc  atcatctgaa  ggtgactctt  cagcagatac  atcagagata  1080
aattctaaac  gtgacagttt  aaacagttct  tcgttgctta  tgaattgctt  cagaaataat  1140
caaagggaag  caaaaagatc  ttggccacc  agatttgacc  tctctgacat  gaagaataca  1200
aagtataact  tggacaagag  gtttggcatg  gattttaag  aaatagaatt  aattggctca  1260
ggtggtttg  gccaggtttt  caaagcaaaa  cacagaattg  acggaagac  ttaagttatt  1320
aaacgtgtta  aatataataa  cgagaaggcg  gagcgtgaag  taagaacatt  ggcaaaactt  1380
gatcatgtaa  atattgttca  ctacaatggc  tgttgggatg  gatttgatta  tgatcctgag  1440
accagtgatg  attctcttga  gagcagtgat  tatgatcctg  agaacagcaa  aaatagtcca  1500
aggtcaaaga  ctaagtgctt  ttcatccaa  atggaattct  gtgataaagg  gaccttgtaa  1560
caatggatgt  aaaaaaag  aggcgagaaa  ctagacaaa  ttttggcttt  ggaactcttt  1620
gaacaaataa  caaaaggggt  ggtattata  cattcaaaaa  aattaattca  tagagattct  1680
aagccaagta  atatacttct  agtagataca  aaacaagtaa  agattggaga  ctttggactt  1740
gtaacctctc  tgaaaaatga  tgaaaagcga  acaaggagta  agggaaactt  gcgatacatg  1800
agccccgaac  agatttcttc  gcaagactat  ggaaggaaag  tggaacctta  cgctttgggg  1860
ctaattcttg  ctgaacttct  tcatgtatgt  gacactgctt  tgaacatc  aaagtttttc  1920
acagacctac  gggatggcat  catctcagat  atatttgata  aaaaagaaaa  aactcttcta  1980
cagaaattac  tctcaaaaga  acctgaggat  cgacctaaac  catctgaaat  actaaggacc  2040
ttgactgtgt  ggaagaaaag  cccagagaaa  aatgaacgac  acacatgtta  gagccctctt  2100
gaaaaagtat  cctgtctctg  atatgcagtt  tctcttaaat  tatctaaaa  ctgctagagg  2160
atatcaatag  atatttacct  tttattttta  tgttttctaa  aattttttac  tatttttact  2220
aatctttctg  cagaaacaga  aaggttttct  tctttttgct  tcaaaaacat  tottacattt  2280
tactttttcc  ttgctcatct  ctattattct  tttttttttt  ttaagacag  agtctcgctc  2340
tgttgccag  gctggagtgc  aatgacacag  tcttggctca  ctgcaactc  ctgctcttgg  2400
gttcaagtga  ttctctgtcc  tcagcctctc  gagtactgtg  atacaggcca  tggcgccacc  2460
accacaactaa  tttttgtgtt  ttaataaag  acaggggttc  accatgttgg  ctgagctggt  2520
ctcaaacctcc  tgacctcaag  taatccactc  gccctggcct  cccaagatgc  tgggattaca  2580
gggatgagcc  accgcgccca  gctctatctc  ttttttctaa  agatggaaaa  accaccccca  2640
aattttcttt  ttatactatt  aatgaatcaa  tcaattcata  tctatttatt  aaattttcac  2700
cgcttttagy  ccaaaaaaat  gtaagatcgt  tctctgtcgc  acatagctta  caagccagct  2760
ggagaatat  ggtactcatt  aaaaaaaaaa  aaaaagtgtat  gtacaacc  2808

```

<210> 106

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002759

<400> 106
tcgtttctctg cctcacatag cttacaagcc agctggagaa atatgttact cattaaaaaa 60

<210> 107
<211> 1678
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002811

<400> 107
aagaagaggg cgcgcgaggg gctgacgaac cggaagaaga ggaactgggc ctgaaagggt 60
acgggtgacc gctactgctg ccggtgtttg cgtgtggcag ggagccagcg ctggcgagcg 120
gggtgtgtcg cgatgccgga gctggcagtg cagaagggtg tgggtccaccc cctggtgctg 180
ctcagtggtg tggatcattt caaccgaatc ggcaagggtg gaaaccagaa gcgtgtgtgt 240
ggtgtgcttt tgggtgcatg gcaaaagaaa gtacttggat tatcgacag ttattgcagt 300
ccttttgatg aagatgacaa agacgattct gtaggtttt tagaccatga tttattggaa 360
aacatgtatg gaatgtttaa gaaagtcaat gccagggaaa gaatagtgtg ctggtaccac 420
acaggcccta aactacacaa gaatgacatt gccatcaacg aactcatgaa aagatactgt 480
cctaattcog tatgtgtcat cattgatgtg aagcgaagg accttagggct gcctacagaa 540
gcgtacattt cagtgggaaga agtccatgat gatggaact caacctcgaa aacatttgaa 600
cacgtgacca gtgaaattgg agcagagaa gctgaggaag ttggagtga acacttgta 660
cgagatatac agacacgac ggtgggcact ctgtcccagc ggatcacaaa ccagggtccat 720
ggtttgaaag gactgaactc caagcttctg gatatcagga gctacctgga aaaagtgcgc 780
acaggcaagc tgcccatcaa ccaccagatc atctaccagc tgcaggacgt cttcaacctg 840
ctgccagatg tcagcttgca ggagttctgc aaggcctttt acctgaagac caatgaccag 900
atggtgtgat tgtacttggc ctgcctgatc cgttccgtg tcgccctgca caacctcatc 960
aacaacaaga ttgccaaccg ggatgcagag aagaagaag ggagagagaa agagagagc 1020
aaaaaggata ggaagagaga caaggagaaa gataaagata aggaagagag tgatgtaag 1080
aaagaggaga aaaaggagaa aaagtataac atgtattaaa tagctttttt aatttgtaaa 1140
ttaaaattctt acaaaactaaa tcagtgtgct gctagagggt tctttttcac ttgacatgct 1200
tattagaagc ctgacccaac aagagctctc tgccctcggg cactcttctg gtgtgtctac 1260
gtggaagtga atggagactg atctcaaatc tgaactgcag ctttcgctgc tgtgagtgtg 1320
ggatatgata gtcagctcag gcttcagatt gatgagaaa aatgaagaga agtcaacaaa 1380
tatatttggtt ctcttcattc atttatctct aaaaaccagga gttgaatttt cctcatcttg 1440
aaagactctt ggggtctgtt tctggtattt tacaaaaatt ctaagtggaa tgcatgaatt 1500
gcattatgtt ctctggtaac acgtagagtt cagacccttc tgaactctgt tgataatacc 1560
acaccatgtt ctggacccat agctctggca tcctcagggg ttgtgtacca gctccatata 1620
ttgtttacct tcaaatgatac aattaaatgg cttgattttt aaaaaaaaaa aaaaaaaa 1678

<210> 108
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002811

<400> 108
aaattgctaa gtggaatgca tgaattgcat tatgttctct ggtaaacagt agagttcaga 60

<210> 109
<211> 846
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002888


```

<400> 109
ccacgtccgg ggtgccgagc caactttcct gcgtccatgc agccccgcgg gcaacggctg 60
cccgtccctt ggtccggggc caggggcccg cgcgccacgg ccccgctgct cgcgctgctg 120
ctgttgctgc ccccggtggc ggcgcccgcg gggtccgggg gccccgacga cccctgggcag 180
cctcaggatg ctgggggtccc gcgcaggctc ctgcagcaga agcgcgcgcg ggcgcttcac 240
ttcttcaact tcgggtccgg ctgcgccagc gcgctgcgag ttgtggccga ggtgcaggag 300
ggccgcgcgt ggattaatcc aaaagaggga tgtaaagtcc acgtggtcct cagcacagag 360
cgctacaacc cagagtccttt acttcaggaa ggtgagggac gtttggggaa atgttctgct 420
cgaagtgttt tcagaatca gaaaccaga ccaaccatca atgtaacttg tacacggctc 480
ctcgagaaaa agaaaaagaca acaagaggat tacctgcttt caaagcaaat gaagcaactg 540
aaaaacccct tggaaatagt cagcatacct gataatcatg gacatattga tccctctctg 600
agactcatct gggatttggc ttctcttggg agctcttacc tgatgtggga aatgacaaca 660
caggtgtcac actactactt ggocagctc actagtgtga gccagtggtt aagaaaaacc 720
tgaaaaataa ctgtgtccoc aagagttaca atcaaaagtg tctccttaga ctgaattcat 780
gtgaactctt aatttcatat caagagtgtt aatcacattt atttcaataa atatgtgagt 840
tcctgc 846

```

```

<210> 110
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002888

```

```

<400> 110
aaagaaaaaa caacaagagg attacctgct ttacaagcaa atgaagcaac tgaaaaaccc 60

```

```

<210> 111
<211> 1054
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_003090

```

```

<400> 111
gaattccgcg ggaggccacg ggccttccac agcgcggggg aacggggaggc tgcaggatgg 60
tcaagctgac ggcggagctg atcagacagg cgcgcgagta caccaacgcg gtgcgcgacc 120
ggagcttgga cctccggggg tataaaattc ccgtcattga aaatctagggt gctacgttag 180
accagtttga tgcatttgat ttttctgaca atgagatcag gaaactggat ggttttccct 240
tgttgagaag actgaaaaca ttgttagtga acaacaacag aatatccgtt atagggtagg 300
gacttgatca ggcctctgccc tgtctgacag aactcattct caaccaataa agtctcgtgg 360
aactgtgtga tctggaccct ctggcatctc tcaaatcgct gaactacota agtatcctaa 420
gaaatccggt aaaccaataa aagcattaca gattgtatgt gattttataa gttccgcaag 480
tcagagtact ggatttccag aaagtgaac taaaagagcg tcaggagaac gagaaaatgt 540
tcaaggggcaa accgggtgca cagcttgcaa aggatattgc caggagaagc aaaaacttta 600
atccaggatg tggtttgcca actgacaaaa agagaggtgg gccatctcca ggggatgtag 660
aagcaatcaa gaatgccata gcaaatgctt caactctggc tgaagtggag aggctgaagg 720
ggttgctgca gtcctggtcag atccctggca gagaacgcag atcaggggccc actgatgatg 780
gtgaagaaga catggaagaa gacacagtca caaacgggtc ctgagcagtg aggcagatgt 840
ataataatag gccctcttgg aacaagtctt gcttttcgaa catggtataa tagccttgtt 900
tgtgttagca aagtggaatc tatcagcatt gtgaaatgc ttaagactgc tgctgataat 960
tttgtaatat aagttttgaa atctaaatgt caattttcta caaattataa aaataaaact 1020
cactctctat gctaaaaaaa aaaaaaagga attc 1054

```

```

<210> 112
<211> 60
<212> DNA
<213> Homo sapiens

```

<300>

<308> NM_003090

<400> 112

taatagcctt gtttgtgtta gcaagtgga atctatcagc attgttgaaa tgcttaagac 60

<210> 113

<211> 2033

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003158

<400> 113

gaattcggg actgagctct tgaagacttg ggtccttggt cgcaggtgga gcgacgggtc 60
 tcactccatt gcccaggcca gagtgcggga tatttgataa gaaacttcag tgaaggccgg 120
 gcgcgggtgct catgcccgta atcccagcat ttctggaggc cgaggcatca tggaccgcatc 180
 taaagaaaac tgcaatttcag gacctgttaa ggctacagct ccagttggag tgccaaaacg 240
 ttgtctcgtg actcagcaat ttcttctgtca gaatccatta cctgtaataa gtggccaggc 300
 tcagcgggtt ttgtgtcctt caaattcttc ccagcgcgtt cctttgcaag cacaaaagct 360
 tgtctccagt cacaaaggcg ttccagaatca gaagcagaag caattgcagg caaccagtggt 420
 acctcatcct gtctccaggc cactgaataa cacccaaaag agcaagcagc ccctgccatc 480
 gcacctgaaa ataactctga ggaggaactg gcatcaaaac agaaaaatga agaatacaaaa 540
 agaggcagtg gctttggaag accttgaaat tgggtgcctt ctgggtaaaag gaaagtttgg 600
 taatgtttat ttggcaagag aaaaagcaaa caagtttatt ctggctctta aagtgttatt 660
 taaagctcag ctggagaaaag ccggagtgga gcatcagctc agaagagaag tagaataaca 720
 gtccacacct cggcatccta atattcttag actgtatggt tatttccatg tatgtaccag 780
 agtctaccta attctggaat atgcaccact tggaaacagt tatagagaac ttcagaaact 840
 ttcaagttt gatgacgaga gaactgctaa ctatataac agaattgcaa atgcctctgc 900
 ttactgtcat tcgaagagag ttattcatag agacattaa cagagaaact tacttcttgg 960
 atcagctgga gagcttaaaa ttgcagattt tgggtggcca gtaactgctc catctctcag 1020
 gaggaccact cctctgtgca ccttggaacta cctgcctcct gaaatgatg aaggtcggat 1080
 gcatgatgag aaggtggatc tctggagcct tggagttctt tgctatgaat ttttagttgg 1140
 gaagcctcct tttgagggcaa acacatacca agagacctac aaaagaatat cacgggttga 1200
 attcacattt cctgactttg taacagaggg agccagggac ctcatctcaa gactgttgaa 1260
 gcataatccc agccagaggc caatgctcag agaagtactt gaacacccctt ggatcacagc 1320
 aaattcatca aaaccatcaa attgccaaaa caaagaatca gctagcaaac agtcttagga 1380
 attcgtgcagg gggagaaatc cttgagccag ggctgccata taacctgaca ggaacatgct 1440
 actgaagttt attttaccat tgactgtctc cctcaatcta caagctaca caagaaatat 1500
 ttgtttttta ctacagaggt gtgccttaac ctccatttc agaaagctcc acatcaataa 1560
 acatgacact ctgaagtga agtagccacg agaattgtgc tactataact ggaacaatat 1620
 ctggaggcaa ggttcgactg cagtcgaaac ttgcctccag attatgaacc agtataagta 1680
 gcacaattct cgtggctact ttcaactcag agtgatcatg ttattgatgt ggagctttct 1740
 gaatagggag gtttaaggc acctgctgag taacacaaat attcttgtg tagcgtttct 1800
 aggaatctgg tctctgtccg gccccggtag gctgtgtgg tttctagtcc tctctaccat 1860
 catctccata tgagagtggt aaaaatggaa cagtgctct acctccattt agggatttgc 1920
 ttgggataca gaagaggcca ttgtctcag agctgttaa ggcttatttt tttaaacatc 1980
 tggagtacata qcattgtgtg aaactttaaa tatgcaggcc gctgtggctc gag 2033

<210> 114

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003158

<400> 114

ttgggtttct agtccctcctt accatcatct ccatatgaga gtgtgaaaa aggaacacgt 60

<210> 115
 <211> 1421
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003258

<400> 115
 acctactcgg ggacggcctt ggagagtact cgggttcgtg aacttcccgg aggcgcaatg 60
 agctgcatta acctgccacc ttgtctgccc ggtcccccga gcaagaccgg gggcgagatc 120
 caggtgatcc tcgggcggat gtctcaggaa aaaagcacag agttgatgag accggtccgt 180
 cgcttcagaa ttgtccagta caagtgcctg gtgatcaagt atgcacaaga cactcgtctac 240
 agcagcagct ttgcacacaa tgaccggaa accatggagg cgtgccgcgc ctgcctgctc 300
 cgagacgtgg cccaggaggc cctgggctg gctgtcatag gcatcgacga gggcgagttt 360
 ttccctgaca tcatggagtt ctgcgaggcc atggccaacg ccgggaagac cgtaatgtgt 420
 gctgcactgg atgggacctt ccagaggaa gcatctgggg ccatctgaa cctggtgccg 480
 ctggccgaga gcgtgggtgaa gctgacggcg gtgtgcatgg agtgcttccg ggaagccgcc 540
 tataccaaga ggctcggcac agagaaggag gtcgaggtga ttgggggagc agacaagtac 600
 cactcgtgt ctggcctctg ctacttcaag aaggcctcag gccagcctgc cgggccggac 660
 aacaagaaga actgcccagt gccaggaaag ccaggggaa cgtggctgc caggaagctc 720
 ttgtcccaac agcagattct gcaatgcagc cctgccaaat gagggacctg caaggcgccg 780
 ccgctccctt cctggcaactg ccgctactg gacgtgcgcc tgcatgtctg ccagccactc 840
 caggaggaa gtcggaggcg tggagggtga ccacacctgt gcctctctgg aactctcctt 900
 ttgtgtgctg ccccacctgc cgcattgctc ctctctcct acccaactgg tgccttaaa 960
 ctctcctctc agctgcgtgg acgatcgccc aggtggagc tggcccccgt ttgtggcctg 1020
 ggatctggca cactccctct ccttgggtg agggacagag ccccacgctg ttgacatcag 1080
 cctgtctctt cccctctgcg gctttcactg ctgagtttct gttctccctg ggaagcctgt 1140
 gccagcacct ttgacccctg gccacactg aggcctagcg ctctctgctt gggatgggct 1200
 cccacccctc cctgaggatg gctggattc agccctctt gttctcttct ggtgtcaaa 1260
 cctctctcac ctctgggtgat ggtttccaca ggaacaacac catctttcac caagatgggt 1320
 ggccacaacc ttgctgggac ttggatccca gggccttatc gtttcaagtg tggagaggcg 1380
 aggtgtccag cctctgctgt agcttatgaa attaactaat t 1421

<210> 116
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003258

<400> 116
 cttctacct cgtgtgatgg ttccacagg aacaacagca tctttcacca agatgggtgg 60

<210> 117
 <211> 913
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003311

<400> 117
 agagccggcg ccgtcaccgc ccgcattgcc gctcccagtc ccgcgctcgg cagcacatga 60
 aatcccccca cgaggtgcta ccgaggggcg agttggagaa gcgcagcgac agcctcttcc 120
 agctatggaa gaagaagcgc ggggtgctca cctccgaccg cctgagcctg ttccccgcca 180
 gcccccgcgc gcgcccacag gagctgcgct tccactccat cctcaagtg gactgctggt 240
 agcgcacggg caagtacgtg tacttcacca tcgtcaccac cgaccacaa gagaatcgact 300
 tccgtcgcgc gggcgagagc tctggaaacg ccggcatcgc ctggcgctc atcgatttcc 360
 agaaccggcg cgcctgcag gactttcgca gccgcaggaa acgcaccgca cccgcccgcac 420

ccgcgcaggag	cgccgtggct	gcgcggcgccg	ccgcaccctc	cgagccctcg	gagccctoca	480
ggccatcccc	gcagcccaaa	cccccgcagc	catgagcccg	ccgcggggcca	tacgctggag	540
gagtcggagcc	gaggctaggga	cgtggccggc	gctctccagc	cctgcagcag	aagaacttcc	600
cgtgcgcggc	gatctcgcct	ccgttgcaag	ggcgccctaa	gttattggac	tatctaatat	660
ctatgtattt	atttcgcgtg	ttctttgtag	tcacatattt	tatatgtctta	atatcttgtt	720
tttgcacac	tttgcccat	gcaaaataat	cacttggcca	gtttgctttt	ctaccatccg	780
gctgtggctc	agtgcagctc	ctcgtgggag	ggtggaggcc	caggaatggg	cgggcaggac	840
accctcatcc	agtctcgcg	ggctggtgtg	aaaggcgctg	ggaaccggct	ttgatgaat	900
aaatgaatcg	tgt	913				

<210> 118

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003311

<400> 118

atttcgctgg tctctttag tagtcacatatt tatagtctta atatcttgtt ttgtcatcac 60

<210> 119

<211> 1723

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003376

<400> 119

tgcggaggc	ttggggcagc	cggttagctc	ggaggtcgtg	gcgctggggg	ctagcaccag	60
cgctctgtcg	ggaggcgccag	cggttaggtg	gaccggtcaag	cggaactcac	ggccaggggcg	120
ctcgggtgctg	gaatttgata	ttcattgatc	cgggtttttat	ccctctctctt	ttttcttaaa	180
caattttttt	taaaactgta	ttgtttctcg	tttttaattta	ttttttgcttg	ccattcccca	240
cttgaatcgg	gccgcagcgt	tggggagatt	gctctacttc	cccaaatcac	tgtggatttt	300
ggaaaccagc	agaaagagga	aagaggtagc	aagagctcca	gagagaagtc	gaggaaagaga	360
gagacggggt	cgagagagagc	gcgcgggcgt	gcgagcagcg	aaagcgacag	gggcaaatgt	420
agtgacctgc	ttttgggggt	gaccgcggga	gcgcggcgctg	agccctcccc	cttgggatcc	480
cgacagctgac	cagtcgcgct	gacggacaga	cagacagaca	ccgccccagc	ccccagctac	540
caactctccc	ccggccggcg	gcggacagtg	gacgcggcgg	cgagccggcg	cgaggggccg	600
gagcccgcg	ccggaggcgg	ggtggagggg	gtcggggctc	gcggcgctcgc	actgaaactt	660
ttcgtccaac	ttctgggctg	ttctcgcttc	ggaggagccg	tggtccgcgc	gggggaagcc	720
gagccgagcg	gagccgcgag	aagtgcctagc	tcggggccgg	aggagcccgca	ggcgaggagag	780
ggggagagg	agaaagagaa	ggaagaggag	agggggccgc	agtggcgact	cggcgctcgg	840
aagccgggct	catggagcgg	tgaggcgccg	gtgtgcgcag	acagtgtctcc	agccgcgcgc	900
gctccccag	ccctggcccc	ggcctcgggc	cggggaggaa	gagtagctcg	ccgagcgccc	960
gaggagagag	ggccgcoccca	cagcccagagc	cggagaggga	gcgcgagccg	gcccggcccc	1020
gctggggcct	ccgaaaccat	gaacttctct	ctgtcttggg	tgcatctggag	ccttgccttt	1080
ctgctctacc	tcacacatgc	caagtggtcc	caaggctgcac	ccatggcaga	agggagaggg	1140
cagaatcatc	acgaagtgg	gaagtctcat	gatgtctatc	agcgacagcta	ctgccatcca	1200
atcgagacc	tgttggacat	ctccaggag	taccctgatg	agatcgagta	catcttcaag	1260
ccatcctctg	tgccctcgat	gcgatcggg	ggctcgtgca	atgacgaggg	cctggaggtg	1320
gtgcaccactg	aggagtcocaa	catcaccatg	cagattatgc	ggatcaaac	tcaccaaggc	1380
cagcacatag	gagagatgag	cttctacag	cacaacaat	gtgaatcgag	accaaagaaa	1440
gatagagcaa	gacaagaaaa	aaaatcag	cgaggaaaag	gaaaggggca	aaaacgaaa	1500
cgcaaaaaat	ccccgtataa	gtcctggagc	gttcctctgt	ggccttgctc	agagcggaga	1560
aagcatttgt	ttgtacaaga	tcgcgacagc	tgtaaatgtg	ctgatcaaaa	caagactcgc	1620
cggtgcaagg	cgagcgagct	tgagttaaac	gaacgtactt	gcagatgtga	caagccggag	1680
cggtgagccg	ggcaggagga	aggagcctcc	ctcagggttt	cgg	1723	

<210> 120
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003376

<400> 120
 ccagcacata ggagagatga gcttctctaca gcacacacaaa tgtgaatgca gaccaaaagaa 60

<210> 121
 <211> 2834
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003406

<400> 121
 gccactccc accgccagct ggaaccctgg ggactacgac gtcctccaaa ccttgcttct 60
 aggagataaa aagaacatcc agtcatggat aaaaatgagc tgggttcagaa ggccaaactg 120
 gccgagcagg ctgagcgata tgatgacatg gcagcctgca tgaagctcgt aactgagcaa 180
 ggagctgaat tatccaatga ggagaggaat cttctctcag ttgcttataa aaatgttgta 240
 ggagcccgta ggtcatcttg gagggtcgct tcaagtattg aacaaaagac ggaaggtgct 300
 gagaaaaaac agcagatggc tcgagaatac agagagaaaa ttgagacgga gctaagagat 360
 atctcgaatg atgtactgtc tcttttggaa aagttcttga tcccacatgc ttoaccaaga 420
 gagagcaaa gcttctctatt gaaaatgaaa ggagattact accgttactt ggctgaggtt 480
 gccctcggtg atgacaaagaa agggattgtc gatcagtcac aacaaagcata ccaagaagct 540
 ttgaaataca gcaaaaaagga aatgcaacca acacatccta tcagactggg ttggccctt 600
 aactctctctg tgttctatta tgagattctg aactccccag agaaaagcctg ctctcttgca 660
 aagacagctt ttgatgaagc cattgctgaa ctgatacat taagtgaaga gctacacaaa 720
 gacagcagcg taataatgca attactgaga gacaacttga cattgtggac atcggtatacc 780
 caaggagacg aagctgaagc agggagaagga ggggaaaatt aacccgctt ccaactttg 840
 tctgctcat tctaaaaatt acacagtaga ccatttgcta tccatgctg cccacaaaa 900
 gtttttctg tacgatttat gacaggttta tgttactctt atttgaattt ctatatttcc 960
 catgtggttt ttaattgttaa tattagggga gttagccag ttaacattta gggagtatc 1020
 tgttttctatc ttgaggtggc caatatgggg atgtggaa ttatacaag ttataagtg 1080
 ttggcatagt acttttggta catttgtgct tcaaaagggc cagtgtaaaa ctgctccat 1140

gtctaagcaa agaaaactgc ctacatactg gtttgtctg gccgggaata aaagggatca 1200
 ttggttccag tcacaggtgt agtaattgtg ggtactttaa gttttggagc acttacaagg 1260
 ctgtgtgata atcataccoc atggatacca catattaaac catgtatc atgtggaaatc 1320
 tcaattgtga cactcttgac tacagctgca gaagtgttcc tttagacaaa gttgtgaccc 1380
 attttactct ggataaaggc agaaaacggtt cacatttccat tatttgaata gttacctgct 1440
 gttagctttc attatttttg ctacactcat ttatttga tttaaatgtt tttaggaacc 1500
 taagaacaaa tgtaaaaagta aagatgcagg aaaaatgaat tgcttggtat tcaattatc 1560
 atgtatatca agcagcagc taaaacaaaa acccatgtat ttaactttt tttaggattt 1620
 ttgcttttgg gatttttttt tttttttttt gatacttgcc ttgctgttaa 1680
 aatagttaac aggggaataa ctbgagatga tggctagctt tgtttaatgt cttatgaaat 1740
 ttctatgaac aatccaagca taattgttaa gaacactgtt attaaatcca gttagtga 1800
 ataaagttt fatgaatgga cttttcaact actttctcta cagottttca tgtaaatttc 1860
 tctgtgttct gaaactcttc taaaggaat ttgacatttt ttgaaattta tctcttattc 1920
 cctcttgcca gctaactggc tcttaccagg tttaaacaca aaattttata taacaaaaat 1980
 actactaata taactactgt ttcatgtccc catgactccc ttcttctc cccacctgta 2040
 aaaaaattgag ttctattttt ttctgggaga ggggggatt gattagaaaa aatgtagt 2100
 tgttccattt aaaaatttgg catatggcat ttcttaactt aggaagccac aaagtcttctg 2160
 gccatcatg acattgggta gcattaactg taagttttgg gcttccaaat cactttttgg 2220
 tttttaagaa ttctctgata ctcttatagc ctgctctcaa ttgtatctt tcttcttct 2280
 tatttgcag gtgcacaaaga ttacttctct gtttagcct tctgtctgt caccacccat 2340
 tcttacttgg tggccatgta cttggaaaaa ggcgcagta tctttctggc tccactcagt 2400

gtctaaggca	cctctgcttc	tttgcttgca	tcccacagac	tatttccctc	atcctattta	2460
ctgcagcaaa	tctctctcta	gttgatgaga	ctgtgtttat	ctccctttaa	aaacctacct	2520
atcctgaatg	gtctgtcatt	gtctgccttt	aaaactcctc	ctctttcttc	ctcctctatt	2580
ctctaataaa	tgatggggct	aagtataacc	caaaagctcc	tttacaaaaa	atttccctag	2640
tactttgcag	aaaaacacaa	acaaaaatgc	cattttaaaa	aagggtgatt	ttttctttta	2700
gaatgtaaag	tcttcaagag	cagggaacaat	gttttctgta	tttctctattg	tgccatgtac	2760
actgttaaat	ctcaataaat	attgatgatg	ggaggcagtg	agtcctgatg	ataagggtga	2820
gaactctgaa	tccc	2834				

<210> 122
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003406

<400> 122
 tttagccttc tgtcttgta ccaaccattc ttacttgggt gccatgtact tggaaaaagg 60

<210> 123
 <211> 1938
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003504

<400> 123
 gatttggcgg gagtcttgac cgcgcgcggg ctcttggtac ctacagcgga ggcgcaggcg 60
 tccggcgccg gtggctatgt tctgtgtcga ttccgcgaaa gaggttctac aggtggtcca 120
 gagccagagg gtccctctct tctgtggcctc ggacgtggat gctctgtgtg cgtgcaagat 180
 ccttcaggcc ttgttccagt gtgaccacgt gcaataatag ctggttccag ttctcgggtg 240
 gcaagaactt gaaactgcac ttcttgagca taaagaacag ttctattatt ttattctcat 300
 aaactgtgga gctaattgtag acctattgga tattcttcaa cctgatgaag acactatatt 360
 ctttgtgtgt gacacccata ggccagtcac gtatcaaacg ataccagat 420
 caaattactc attaaacaag atgatgacct tgaagttccc gccatgaag acatcttcag 480
 ggatgaagag gaggatgaag agcattcagg aaatgacagt gatgggtcag agcctcttga 540
 gaagcgcaac cggttagaag aggagatagt ggagcaaac atgcggagga ggcagcgccg 600
 agagtgggag gcccgagaa gagacatcct ctttgaactac gagcagatg aatatcatgg 660
 gacatcgtca gccatggtga tgtttgagct ggcttggatg ctgtccaaag acctgaatga 720
 catgctgtgt tgggccatcg ttggaactaac agaccagtgg gtgcaagaca agtcaactca 780
 aatgaataac gtgactgatg ttgtgtctct gcagcgccac gtttccgcc accaaccacg 840
 gaaacgaggt gaggagaaca caactctcgt gctctgcaca cggatctctc ttgagtatga 900
 cctccgctgt gtgctctacc agcactggtc gactccatgac agcctgtgca acaccagcta 960
 taccgcagcc aggttcaagc tgtggtctgt gcatggacag aagcggctcc aggaattctc 1020
 tgcagacatg ggtcttcccc tgaagcaggt gaaagcagaag ttccaggcca ggcacatctc 1080
 cttgaaggag aatttgcggg aaatgatgga agagtctgca aataaatttg ggatgaagga 1140
 catgcgcgtg cagactttca gcaattcatt ttgggtcaag cacaagtttc tggccagcga 1200
 cgtggtctct gccacactgt ctttgatgga gagccccag aaggatggct cagggaaga 1260
 tcaacttcat caggctctgg acagcctctc caggagtaac ctgacaagc tgtaccaagt 1320
 cctggaactc gccaaagaag agctgcgagc caccagcag accattgcca gctgcctttg 1380
 caccacctc gtcatctccc aggggccttt cctgtaactc tctctcatg tctaccactc 1440
 agatgtcatg ctgttctcta ggccgcatc cctaaagctg ctacgaaac acctgctcaa 1500
 gtcccttgtg tggctgacaa agaaccggcg ctgcaaaact ctgcccctgg tgatgctgc 1560
 cccctcgagc atggagcatg gcacagtgac cgtggtgggc atccccccag agaccgacag 1620
 ctggacagag aagaactcttt ttgggagggc gcttgagaag gcaagcgaaa gcaccagctc 1680
 ccgatgctg cacaacattt ttgacctctc agtaattgag ctgaaagctg aggatcgag 1740
 caagtctctg gacgcactta ttccctctc gtacctagaa ttgtattctc ccagaaagac 1800
 ctctctattt atgtaactgg ctctcaattg gatgtgaagt tatggacatg atttgagatg 1860
 tagaagcat tttttattaa ataaaatgct tattttaggc tccgtcccca aaaaaaaaaa 1920

aaaaaaaaa aaaaaaaaa 1938

<210> 124
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003504

<400> 124
caagtttctg gacgcactta tttccctcct gtccctaggaa ttgtattctt ccagaatgac 60

<210> 125
<211> 2346

<212> DNA
<213> Homo sapiens

<300>
<308> NM_003600

<400> 125
acaagcgacg ctgcgtcgag cgcagggcaa tcggtttct agctagaggg tttactcct 60
atttaaaag aagaaccttt gaattctaac ggctgagctc ttggaagact tgggtccctg 120
ggctcgacgt ggagcgacac ggttgggtag accgtggggg atatctcagt ggcggacgag 180
gacggcgggg acaagggcg gctggtcgga gtggcgagac gtcaagtcct cgtcggttc 240
ctccgtccct gagtgtcctt ggcgtgcct tgtgccgcc cagcgccctt gcactccgtc 300
ctgggcacgg agggccctcg taggtactgt ctgttactt attacagcta gaggcatcat 360
ggaccgatct aaagaaaact gcatttcagg acctgttaag gctacagctc cagttggagg 420
tccaaaacgt gttctcgtga ctacgaatt tccttgtcag aatccattac ctgtaaatag 480
tgccacggct cagcgggtct tgtgtccttc aaattcttcc cagcgcttc ctgtgcaagc 540
acaaagcgtt gtctccagtc acaagcgggt tcagaatcag aagcagaagc aattgcaggc 600
aaccagtgta cctcatcctg tctccaggcc actgaataac acccaaaaga gcaagcagcc 660
cctgccatcg gcacctgaaa ataactctga ggaggaaact gcactcaaac agaaaaatga 720
agaatcaaaa aagaggcagt gggctttgga agactttgaa attggtcgcc ctctgggtaa 780
aggaaagttt ggtaatgttt atttggcaag agaaaagcaa agcaagttta ttctggctct 840
taaagtgtta tttaaagctc agctggagaa agccggagtg gagcatcagc tcagaagaga 900
agtagaataa cagtcaccac ttccgcatcc taatattctt agactgtatg gttatttcca 960
tgatgtacc agatgtctacc taattctgga atatgcacca cttggaacag tttatagaga 1020
acttcagaaa ctttcaaaagt ttgatgagca gagaactgct acttatataa cagaattggc 1080
aaatgccctt ctttactgtc attcgaagag agttattcat agagacatta agccagagaa 1140
cttactctctt ggatcagctg gagagcttaa aattgcagat ttgggtgggt cagtacatgc 1200
tccatcttcc aggagagcca ctctctgtgg caccctggac tacctgcccc ctgaaatgat 1260
tgaaggtcgt atcgatgatg agaaggtgga tctctggagc cttggagttc ttgtctatga 1320
atttttagtt gggaagcctc ctcttggagg aaacacatac caagagacct acaaaagaa 1380
atcacggggt gaattcacat tccctgaact tgaacagag ggagccaggg actcatttc 1440
aagactgttg aagcataatc ccagccagag gccaatgctc agagaagtagc ttgaacaccc 1500
ctggatcaca gcaaatctcat gggggagaaa tcttggagcc aggggtgcc aataacctga 1560
acagctcttag gaatcgtgca ttattttacc attgaactgt gccctcaatc tagaagctca 1620
caggaacatg ctactgaagt ttattttacc attgaactgt gccctcaatc tagaagctca 1680
cacaagaatc atttgtttta ctacgacagt gtgccttaac ctccctattc agaaagctcc 1740
acatacaaaa acatgacact ctgaagttaa agtagccaac agaattgtgc tacttatact 1800
ggttcataat ctgagagcaca ggttcgactg cagccgcccc gtcagcctgt gctagcgctg 1860
gtgtctttac agtggagcaca tccagagcct ggctgtgggg aaagtgaacca gttgccctgt 1920
accocgatca gtttaaggagc tgtgcaataa ccttctcagt acctgagtga gtgtgtaact 1980
tattgggtgt gcgaagcctg tgaagctgt atgtgattct ttttaagtat 2040
gaaaataaag atatatgtac agacttgtat ttttctctg gtggcattcc tttaggaatg 2100
ctgtgtgtct gtccggcacc ccggtaggcc tgattgggtt tctagctctc cttaaacact 2160
tatctcccat atgagagtggt gaaaaatagg aacacgtgct ctacctccat tttagggatt 2220
gcttgggata cagaagaggc catgtgtctc agagctgtta agggcttatt tttttaaacc 2280

attggagtca tagcatgtgt gtaaaacttta aatgtgcaaa taaataagta tctatgtcta 2340
 aaaaaa 2346

<210> 126
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_003600

<400> 126
 agagtggtgaa aaataggaac acgtgctcta cctccattta gggatttgct·tgggatacag 60

<210> 127
 <211> 853
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_003641

<400> 127
 ctagtctctga cttcacttct gatgaggaag cctctctcct tagccttcag cctttcctcc 60
 caccctgcc aagttaattt gatcctcaag aagttaaacc acacotcatt ggtccctggc 120
 taattcacca atttacaacac agcaggaaat agaaacttaa gagaaataca cacttctgag 180
 aaactgaaac gacaggggaa aggaggtctc actgagcacc gtccocagcat ccggacacca 240
 agcgggccct tgcctccacg cagaaaacca cacttctcaa accctcaetc aacacttctct 300
 tccccaaagc cagaagatgc acaaggagga acatgaggtg gctgtgctgg gggcaccccc 360
 cagcacatc cttccaaagt ccaaccgtgat caacatccac agcgagacct ccgtgccccga 420
 ccatgtctgc tggctccctgt tcaacacctt ctctctgaac tgggtgctgtc tgggcttcat 480
 agcattcgcc tactccgtga agtctaggga caggaagatg gttggcgacg tgaccggggc 540
 ccaggcctat gctccaccg ccaagtgcct gaacatctgg gccctgattc tgggcatcct 600
 catgacattt ggattcatcc tgtcactggt attcggtctc gtgacagctt accatattat 660
 gttacagata atacaggaaa aacgggggta ctagtaccgg cccatagcct gcaacctttg 720
 cactccactg tgcaatgctg gccctgcacg ctggggctgt tgccccctgcc ccttgggtcc 780
 tgccccctaga tacagcagtt tatacccaca cacctgctcta cagtgtcatt caataaagtg 840
 acgtgcttg tga 853

<210> 128
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_003641

<400> 128
 attatgttac agataatata ggaaaaacgg ggttactagt agcgcgccat agcctgcaac 60

<210> 129
 <211> 1280
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_003756

<400> 129

gaaagatggc	gtcccgaag	gaaggtagcg	gctctactgc	cacctcttcc	agctccaccg	60
ccggcgccagc	agggaaagc	aaaggcaaa	gcggctcggg	agattcagcc	gtgaagcaag	120
tgacagtaga	tgcccttggt	gtattaaaga	taatcaaa	ttatcaagaa	gaaggacaag	180
gaactgaagt	tggtcaagga	gtgcttttgg	gtctggttgt	agaagatcgg	cttgaaatta	240
ccaactgctt	tcctttccct	cagcacacag	aggatgatgc	tgactttgat	gaagtccaat	300
atcagatgga	aatgatgcgg	agccttcgcc	atgtaaacat	tgatcatctt	cagctgggct	360
ggtatcagtc	cacatactat	ggctcattcg	ttaccggggc	actcctggac	tctcagttta	420
gttaccagca	tgccattgaa	gaatctgtcg	ttctcattta	tgatccata	aaaaactgcc	480
aggagtctct	ctcactaaag	gcatacagac	tgactcctaa	actgatggaa	gtttgtaaag	540
aaaagatttt	ttccctgaa	gcattgaaaa	aagcaaatat	cacctttgag	tacatgtttg	600
aaagaagtgc	gattgttaatt	aaaaattcac	atctgatcaa	tgctcctaat	tgggaaacttg	660
aaaagaagtc	agctgttgca	gataaacatg	aattgtccag	ccttgccagc	agcaatcatt	720
tggggaagaa	tctacagttg	ctgatggaca	gagtgatga	aatgagccaa	gatatagtta	780
aatacaacac	aaactgttgc	aatactagta	aacaacagca	gcagaaacat	cagtatcagc	840
agcgtgccca	gcaggagaat	atgcagcgcc	agagccgagg	agaacccccc	ctccctgagg	900
aggacctgtc	caaaactctt	aaaccaccac	agccgctgcg	caggatggac	tcgctgtcca	960
ttgcaggcca	gataaacact	tactgcccaga	acataagga	gttcaactgcc	caaaacttag	1020
gcaagctctt	catggcccag	gctcttcaag	aatacaacaa	ctaagaaaa	gaagttttcca	1080
gaaaagaagt	taacatgaac	tcttgaaagtc	acaccagggc	aactcttgga	agaaatata	1140
ttgcataatt	aaaagcacag	aggatttctt	tagtgtcatt	gccgattttg	gctataacag	1200
tgtctttcta	gcataataaa	aataaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	1260
aaaaaaaaa	aaaaaaaaa	1280				

<210> 130

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003756

<400> 130

tgagccaaga tatagttaaa tacaacacat acatgaggaa tactagtaaa caacagcagc 60

<210> 131

<211> 839

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003832

<400> 131

aaagccacag	ctccctggct	ggcgtagct	aaagtggctg	ttgggtgtcc	gcaggcttet	60
gcctggccgc	cgccgcctat	aagctaccag	gaggagcttt	acgaactccc	gtcctgcggg	120
aagtggcggg	cacgatcgca	aggtagcgca	gaagcttctc	aatggccagc	gcagctgcoa	180
gcccgcggcg	cgcaactgcc	tcacctgagc	ctgggaggaa	aattcttcca	aggatgatct	240
cccaactcaga	gctgaggaa	cttttctact	cagcagatgc	tgtgtgtttt	gatgttgaca	300
gcacggtcat	cagtgaaaga	ggaatcggat	gctttcattg	gatttggagg	aaatgtgatc	360
aggcaacaag	tcaaggataa	cgccaaatgg	tatatcactg	attttgtaga	gctgctggga	420
gaaccggaag	aataacatcc	attgtcatac	agctccaaac	aacttcagat	gaatttttac	480
aaagtacaca	gattgatact	gtttgtttac	aattgcctat	tacaacttgc	tataaaaaat	540
tggtacagat	gatctgcact	gtcaagtaaa	ctacagttag	gaatcctcaa	agatttggttt	600
gtttgttttt	aactgttagtt	ccagtatata	atgatcacta	tcgatttcc	ggagagtttt	660
gtaactctgaa	ttcttttagtt	atattcctag	ctataattca	tacaagtggt	tttaagagtg	720
gagagtgcaat	taaacacctt	tactcttagg	aatatagatt	cggcagcctt	cagtgaatat	780
tggttttttt	ccctttggta	tgtcaataaa	agtttatcca	tgtgtcagaa	aaaaaaaaa	839

<210> 132

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003832

<400> 132

gaagaaggaa tcggatgctt tcatgtgatt tggaggaaat gtgatcaggc aacaagtcaa 60

<210> 133

<211> 3128

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003981

<400> 133

gcttcgcccc	gtggcgcggt	tgaaaathtt	goggggctca	acggctcgcg	gagcggctac	60
gcggagtgac	atcgccgggt	tttgcgggtg	gttgttgctc	tcggggccgt	gtggagtagg	120
tcgtgacctg	gactcacggc	tgcttggagc	gtccgccatg	aggagaagtg	aggtgctctg	180
ggaggagctcc	atagtatgtc	tgcagaaagc	cctaatac	cttcgggaaa	tatgggagct	240
aattgggatt	ccagaggacc	agcggttaca	aagaactgag	gtggtaaa	agcatatcaa	300
ggaactcctg	gatattgatg	ttgctgaaga	ggaagccctg	aaggaaagac	tcatcaaaag	360
catatccgtc	tgtcagaaga	agctgaacac	tctgtgcagc	gagttacatg	ttgagccatt	420
tcaggagaaga	ggagagacga	ccatcttgca	actagaaaaa	gatttgcgca	cccaagtgtga	480
attgatgcga	aaacagaaaa	aggagagaaa	acaggaaactg	aagctacttc	aagagcaaga	540
tcagaactgt	tgcgaaatct	tttgtatgcc	ccactatgat	attgacagtg	cctcagtgcc	600
cagcttagaa	gagctgaacc	agttcaaggca	acatgtgaca	actttgaggg	aaacaaaggc	660
ttctaggcgt	gaggagtttg	tcagatataa	gagacagatc	atactgtgta	tggagaattt	720
agaccacacc	ccagacacaa	gctttgaaag	agatgtgggt	tgtgaagacg	aagatgcctt	780
ttgtttgtct	ttggagaata	tgtcaacact	acaaaagtgt	ctacggcagc	tggaagtgtca	840
gaaatcacaa	aatgaagcag	tgtgtgaggg	gctgcgtact	caaataccag	agctctggga	900
caggttgtcaa	atacctgaag	aagaaagaga	agctgtggcc	accattatgt	ctgggtcaaa	960
ggccaaaggt	cggaaagcgc	tgcacttaga	agtggatcgg	ttggaagaac	tgaataatgtca	1020
aaacatgaag	aaagtgtatt	aggaacttcg	attggagctg	gttcagtact	gggaccagtg	1080
cttttatagc	caggagcaga	gacaagcttt	tgccctcttc	tgtgctgagg	actacacaga	1140
aagtctgtct	cagctccacg	atgctgagat	tgtgcggtta	aaaaactact	attgaagtcca	1200
caaggaaact	tttgaagtg	tcagagaagt	ggaagaagcc	tggaagcttt	tcttagagttt	1260
tcagagaaaa	gcttcagatc	caaatcgatt	acaaaccgga	ggaggaatc	tctcaaaaga	1320
agaaaaacaa	cgagcccaagc	tcagaaaaat	gctgcccaag	ctggaagaag	agttgaaggc	1380
acgaattgaa	tgttgggaac	aggaacattc	aaaggcattt	atgggtgaat	ggcagaattt	1440
catggtgat	gtggcagaac	aatgggagat	gcactgattg	gagaagaaga	gagccaagca	1500
ggaagaccaa	ctgaagaaca	aaaaacagac	agagacagag	atgctgtatg	cgagcgctcc	1560
tcgaacacct	agcaagcggc	gaggactggc	tcacaaatca	ccgggcaag	cagctaagct	1620
gaacactacc	accatgtcca	atgtcacggc	taactagtgc	attcggccta	tctttggagg	1680
gacagctcac	cactccccgc	tgtctcgact	tcctcctctg	ggcagcaagc	cagtcgctgc	1740
ttccacctgt	tcagggaaga	aaacaccocg	tactggcgca	catggagcca	acaaggagaa	1800
cctggagctc	accgcgacga	tcctgagtg	tggttacctc	ggctcgggcc	ccctccagcg	1860
caacttcagc	attaattctg	ttgcagacac	ctattctgag	tttgcaagg	atccgtccctc	1920
ctctgacagt	tcacactgtg	ggcttcacgc	agaacttcca	aggcttcca	aatctgatgc	1980
tacttctgga	tctctcaatt	caaccaacat	ccagtcccta	gaagccctga	tcagtcacac	2040
agctgtggct	tcctgtgcct	agactggacc	taattatagt	ggggtgactt	tagtttttct	2100
tcagcttagg	cgtgcttgaa	accttggcca	ggttccatga	ccatgggctc	aacttaaaag	2160
tgtaagttag	tggttacagtt	gaaagcccat	cataggttta	gtggtcctag	gagacttggt	2220
tttgacttat	atacatgaaa	agtttatggc	aagaagtgtca	aattttagta	tatggggcct	2280
gacttctcta	ccacataatt	ctacttgctg	aagcatgac	aaagcttgtt	ttatttcacc	2340
actgtaggaa	aattgattgac	tatgccatc	cctgggggta	attttggcat	gtatacctgt	2400
aactagtaat	taacatcttt	tttgttttag	catgttcaag	taatgtctag	gctatcatag	2460
ctttgtctct	acctgaagcc	tgtgtccccc	cacacaggac	agccttcctc	ctgaagagaa	2520

tgtctttgtg	tgtccgaagt	tgagatggcc	tgcctactg	ccaaagaggt	gacaggaagg	2580
ctgggagcag	ctttgttaaa	ttgtgttcag	ttctgttaca	cagtgcatgt	ccctttgttg	2640
ggggtatgca	ttgatgaaca	cacatgcttg	tcggaaacgt	ttctcggcgt	ttgtcccttg	2700
gctctcatct	cccccatc	tgtgctact	ttgctgagt	tcttctacc	cgcagtgctg	2760
cagccacatt	gggagtcgt	ttgttccaat	gggttgagct	gtctttgtcg	tggagatctg	2820
gaactttgca	catgtcacta	ctggggaggt	gttctgtct	tagcttccac	gatgaggcgc	2880
ctcttttacc	tatcctctca	atcactactc	ttcttgaagc	actattattt	attcttccgc	2940
tgtctgctgt	cagcagctgt	actgtcaaca	tagtgaat	ggttctcaaa	agatttaccag	3000
tgtggactgt	gtgttagcca	cgtgtttac	tcatacagta	cgtgtcctgt	ttttaaaaata	3060
tacaattatt	cttaaaaaata	aattaaaatc	tgtatactta	catttcaaaa	agaaaaaaa	3120
aaaaaaa	3128					

<210> 134

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003981

<400> 134

tgccagcagta	ctactgtcaa	catagtgtaa	atggtttctca	aaagcttacc	agtggtgact	60
-------------	------------	------------	-------------	------------	------------	----

<210> 135

<211> 1816

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004029

<400> 135

ggcaccagc	gtccggcctg	cgccttcccg	ccaggcctgg	acactggctc	aacacctgtg	60
actttcatgt	tgcgcgcggg	ccacacctgc	agtcacacct	gtagccccct	ctgcgaagag	120
atccataccg	aggcagcgtc	gggtggctaca	agccctcagt	ccacacctgt	ggacacctgt	180
gaacacctggc	cacacgacct	gtggccgcgg	cctggcgctt	gctgcgacag	gagcctttac	240
ctccccctgt	ataaacacct	accgccacct	aactgccctt	gcagaaggag	caatggcctt	300
ggctcctgag	aggcgagccc	cacgcgtgct	gttcggagag	tggctccttg	gagagatcag	360
cagcggctgc	tatgaggggc	tgcagtgctg	ggcagggccc	cgcacctgtt	tcgcgctgcc	420
ctggaagcac	ttcgcgcgca	aggacctgag	caggagccgac	gcgcgcatct	tcaagcctgt	480
ggctatggcc	cgccggcaggt	ggccgcctag	cagcagggga	ggtggcccg	ccccgaggc	540
tgaactgcg	gagcgcgcg	gctggaaaac	caacttccgc	tgcgcactgc	gcagcacgcg	600
tcgcttcgtg	atgctgcggg	ataactcggg	ggaccggccc	gaaccgcaca	aggtgtacgc	660
gctcacccgg	gagctgtgct	ggcgagaagg	ccagggcacg	gaccgacgtg	aggcagaagg	720
ccccgagct	gtccccacac	cacagggtgg	gcccccaggg	ccattcttgg	cacacacaca	780
tgtctgactc	caagcccacg	gcccccctcc	tgcctccagc	ggtgacaagg	gggacctctt	840
gctccaggca	gtgcaacaga	gctgcctggc	agaccatctg	ctgacacgct	catggggggc	900
agatccagtc	ccaacgaagg	ctctctggag	gggacaaaga	gggcttcccc	tgactggggc	960
ctgtgctgga	ggcgagggcg	cggcccagca	gtccccgcac	caggcgagag	cgtaactgtc	1020
accctcccaca	agcgctctca	cgcggttgca	agagcccacg	ccaggggcgc	tggacgtgac	1080
catcatgtac	aaggggccga	cgtgtgtgca	gaaggtggtg	ggacaccgca	gctgcacgtt	1140
ctctatcggc	ccccagacc	cagctgtccg	ggccacagac	ccccagcagg	tagcatcccc	1200
cagccctgcc	gagctcccg	accgaagaca	gctgcgctac	acggaggaac	tgtctgcggca	1260
ctgtgcccct	gggttgacc	tggagctctg	ggggccacag	ctgtggggcc	ggcgcatggg	1320
caagtgcagg	gtgtactggg	aggtggcgcg	acccccaggc	tcggccagcc	ctccaccccc	1380
agcctgcctg	ctgcctcgga	actgtgacac	ccccattctt	gaattcagag	tccttctcca	1440
agagctgggt	gaattccggg	cacggcagcg	ccgtggctcc	ccacgctata	ccattctacct	1500
gggcttcggg	caggacctgt	cagctgggag	gcccaaggag	aagagcctgg	tcctgtgtgaa	1560
gctggaaccc	tgctctgtcc	gaagtgcacct	agagggcacg	cagcgtgagg	gtgtgtcttc	1620
ctcgatagc	agcagctcca	gcctctgctt	gtccagcgcc	aaacgctctt	atgacagcat	1680
cgagtgtctt	cttatggagc	tggagcagcc	cgcctagaac	ccagctctaat	gagaaactcca	1740

gaaagctgga gcagcccaacc tagagctggc cgcggccgcc cagtctaata aaaagaactc 1800
cagaacaaaa aaaaaa 1816

<210> 136
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004029

<400> 136
agcagcccac cttagagctgg ccgcgccgcg ccagtctaata aaaagaact ccagaacaaa 60

<210> 137
<211> 2121
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004203

<400> 137
tggaaatTTTT ggcgcgagca gctccgcgcg cgttcacggg ccgttccccc tcacgggagt 60
cctccgcccg ggcgtccgga acagtcgacg gcagactccg gcccgctgag ccacccgagg 120
ggtcocctgg cctccgcgga cccggaatct gggccctcgc ggacccgcgc cccgcccatg 180
cgccccaggg cttccccaca cccacggagt gaagtcaagg gcggccctgc ctgggaggaa 240
cttaccgtct accgggaaaag gtggccagca ggtgtgtcgg gctcgtgag agggtagagc 300
gagacggccc gatcgccagc ggccccgaaa gctgcggagg tcacccccgc ctggccctag 360
ctcaaggaca cctcggattc acgtgggagc cctgctcct gcctccccg tcaccacct 420
gaggctgttg ggccaggcca gtcatgctag aaocgctcct tgcactggcc atgcccatgc 480
ccacggaggg caccocgcca cctctgagtg gcacccccat cccagtccca gctacttcc 540
gccacgcaga acctggattc tccctcaaga ggccccgggg gctcagccgg agcctccca 600
ctccgcccc tgccaaagggc agcattccca tcagccgcct ctccctcct cggaccccag 660
gctggcacca gctgcagccc cggcgggtgt cattccgggg cgaggcctca gagactctgc 720
agagccctgg gtatgacca agccggccag agtccttctt ccagcagagc ttccagagc 780
tcagccgctt gggccatggc tctacggag aggtcttcaa ggtgcgctcc aaggaggacg 840
cgcgctcta tgcggtaaag cgttcctatg caccattccg ggcccccaag gccggggccc 900
cgaagtggcg cgaggtgggc agccacgaga aggtggggca gcacccattg tgcgtgcggc 960
tggagcagcg ctgggaggag gcggcctcc tgtaactgca gacggagctg tcggcgccca 1020
gcctgcagca acactgtgag gcctgggggt ccagcctgcc tgaggccagc gtctgggggt 1080
acctgcggga ccgcgtgctt gccctggccc atctgcacag ccaggggcgt gtgcacctgc 1140
atgtcaagcc tgccaacatc ttctggggc cccggggccg ctgcaagctg ggtgacttgc 1200
gactgtgctg ggaagtgggt acagcaggag ctgggtgaggt ccaggaggga gacccccgct 1260
acatggcccc cgagctgctg cagggtcctt atgggacagc agcgatgtg ttcagtgtgc 1320
gcctcaccat cctggaagtg gcatgcaaaa tggagctgcc ccacggtggg gagggtctgc 1380
agcagctgcg ccagggtctac ctgcocctgc agttcactgc cggctctgct tccagagcgc 1440
gttctgtctt tgtcatgatg ctggagccag accccaagct gcggggccag gccagggccc 1500
tgtggcaact gctctgtgtg aggcagccgc gggcctgggg tgtgctgtgg tgcattggcag 1560
cggaggccct gagccgaggg tgggccctgt ggcaggccct gcttgcctgt cctctgtggc 1620
ctggtcatgg gctgcctcac cctgccagct gctacagcg cctgcctcag aactgggatg 1680
gcctggctc accacccctgc agtctgtcct tggacagcag cctctccagc aactgggatg 1740
acgcagcctc agggccttca cctccctcct aggtctcct ggccecgact gtggggagca 1800
cctccacccc ccgggacagg tgcaacccca gggatgcctt ggaactaaat gacatcaact 1860
cagagcctcc tctggggctcc ttccctcctt ttgagccttg gaactccttc agcctgtttg 1920
aggcaacctc agaccaaccc tgagccccag actctgcctc tgcaacttta accttttatc 1980
ctgtgtctct cccgtgcctc ttgaaagctg gggccctcgc ggaactccca tggctctctc 2040
tgctgtgccg tgtctaataa aaagtatttg aaacttggga gcacccaagc ttgctcatgt 2100
ggcaaaaaaa aaaaaaaaaa a 2121

<210> 138

<211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004203

<400> 138
 ctggcgcgtgt ctaataaaaa gtatttgaac ctgtggagca cccaagcttg ctcattgtgc 60

<210> 139
 <211> 1982
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004207

<400> 139
 ggccagaggg ggccgtgagg ggcccagcgg cggcaggtga ggccgaacca accctcctgg 60
 ccatgggagg ggccgtgggt gacgagggcc ccacaggcgt caaggccctt gacggcggt 120
 ggggctgggg cgtgctcttc ggctgtttcg tcatcactgg cttctctac gcttcccca 180
 aggcgtcag gtctctcttc aaggagctca tacaggagt tgggatcgg tacagcgaca 240
 cagcctggat ctctctcatc ctgctggcca tgcctacgg gacaggctcg cttcgcagt 300
 tgtgctgaa ccgctttggc tgccggcccg tcatgcttg ggggggtctc tttgctgcg 360
 tgggcctggt ggctgcttcc ttttgcggga gcatcatcca ggtctacctc accactggg 420
 tcatcacggg gtgggtttg gcactcaact tccagccctc gctcatcatg ctgaacggct 480
 acttcagcaa gggcgccccc atggccaaag ggtggcggc agcaggtagc cctgtcttcc 540
 tctgtgccct gaggccgctg gggcagctcg tgcaggaccg ctacggctgg cggggcggt 600
 tccatctctt gggcgccctg ctgctcaact gctgcgtgtg tgcgcactc atgaggcccc 660
 tgggtgtcac gggccagccg ggctcggggc cgcgcgaccc ctcccgccgc ctgctagacc 720
 tgagcgtctt ccgggacccg ggctttgtgc ttacgcctg ggccgctcg gtcattgtgc 780
 tgggctctt cgtcccgccc gtgtttcgtg tgagctacgc caaggacctg ggctgccc 840
 acaccaaggg cgccttctcg ctaccatcc tgggtctcat tgacatctt cgcgcggccg 900
 ccgcggtctt cgtggcgggg ctgtgggaag tgcggcccta ctccgtctac ctcttcagt 960
 tctccatgtt cttcaacggc ctccggacc tggcgggct tacggcgggc gactacggcg 1020
 gctcgtggt cttctgcatc ttctttggca tctctacgg catggtggg gccctcgat 1080
 tcgaggtgct catggccatc gtgggcaccc acaagttctc cagtgcatt ggctgggtgc 1140
 tgcattgga ggccgtggcc gtgctcgtcg ggcctcttc gggaggcaaa ctctctggtg 1200
 cgaccacgt ctacatgtac gtgttcatcc tggcgggggc cgaggtgctc acctctccc 1260
 tgattttgct gctgggcaac ttcttctgca ttaggaaag gcccaagag ccacagcgtg 1320
 aggtggcgcc cggcaggagg gagaagctcc acaagcctcc tgcagactc ggggtggact 1380
 tgcgggaggt ggaacatttc ctgaaggctg agcctgagaa aaacggggag gtggttcaca 1440
 cccgggaaac aagtgtctga gtggctgggg ggggcccga ggcacaggga ggaagtacag 1500
 aagcgcgcaa cgtttgctat ttttttaca aactggactg gctcaggcag gctcacggct 1560
 gggctccagg tgcgggcccc cgggatcgt gcccgatcag tgttttgagg gggaaagtgg 1620
 cgggtggga accgtgtcat tccagagtgg atctcggtg aagccaagcc gcaaggttac 1680
 aagcatcct caccaggggc ccgcctgct gctccaggt ggctcgggc cactgctatg 1740
 ctcaagacc ttgaaaccca tgcctcgaga caactgact ttaatgggag ggtgggtggg 1800
 ccgcaacag gctgcccagg caggtcctgc gtggggcctc ctccagcccg tcttaccctg 1860
 ggctcacatg ggccctgtgc ccaccctct tgaagtgtt ggggacagct ctttccacc 1920
 ctggaaggtt gaaataaacc tgcgtgtgg tggagtgttc tctgcccga ttcaaaaa 1980
 tt 1982

<210> 140
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004207

<400> 140
 cctcttgagtg gtcttgggga cagctctttc cacccttgga agatggaat aaacctgcgt 60

<210> 141
 <211> 2054
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004209

<400> 141
 cgggaggcgg cagcggtcgc agcgttggtg gcacatcagc cagcatcagc ggcagcgcca 60
 gcggcctcgg gcggggccgg ccggacggac aggcggacag aaggcgccag gggcgcgct 120
 ccgcgccggg ccggccatgg agggcgccct cctcgccggc ggccgcgagc gggccgcct 180
 ggcacccgtg agcttgcgc ggccggccca gacctgctc cgggtcgct cctgggtgtt 240
 ctccatcgcc gtcttcgggc ccatcgtcaa cgagggtac gtgaacacgc acagcgccgc 300
 cgagctcgcc tgcgtgttca acgggaacgc gggcgccctc cgcttcggcg tcgcgtggg 360
 cctcgagacc ttctcgctt gcgcgcctt cctgctgctc gatgtcgct tccagcaaat 420
 cagcagcgtc cgcgaccgcc ggccgcgggt gttgctggac ctgggctctt caggactctg 480
 gtctcttcct tggttcgtgg gcttctgctt cctcaccaat cagtggcagc gcacggcgcc 540
 agggcgccgc acgacgcagg cgggggacgc ggcgcgggcc gcatcgctt tcagcttctt 600
 ctccatctc agctgggtgg cgctcacctg gaaggccctc cagcgggttc cctggggcac 660
 gcacatgtca ctcttcgcca ccgaacagct gagcaccggg gcgagccagg cctaccocgg 720
 ctatccggtg ggcagcgcg gcaggggac tggaaggcac cgagacctac cagagcccgcc 780
 gacctcgacc accagcccca aagggtacca ggtgcccgcc tactagcgcc tggcaggcac 840
 agaccaggcg tccaaggcca ccccaccaac gcaggcccca ggggtctccgg gacctccctt 900
 ggggtccttc agctcagtcgc cgcggacaga gtagggtggc gcttgcgcc atccggggcc 960
 aagagggggt ggaaccgcgt gtctgggctg ccctgccaa gttcccccag tccctcagca 1020
 cctggcccca ggaactgaggt cctgagaagg ggaatagcat gctgtcccta tgtgtcccta 1080
 gcctggaatg gactggcctg gggaaggctt tccctcttg ggccacacct gctcactctg 1140
 ggggtggggg tccagctgcc ctctacgac aggtgcagg gcgcccagg acaaggcggg 1200
 ggcaagggaag agacaccacc ctgcgcccaa gactggggat cctggccact gtcccatcc 1260
 catgtccctg tgggtagtga ctgtctgctt tctgtcatgg tgggtcgctc cgtccggagc 1320
 cactctccac ttctctcac aggtgctag aacagccag cctgtcagt gttgtgatca 1380
 tggtcagtc ttccgggttc acctcctagt actccacaag ctgctcctc ctctgtggcc 1440
 ccggccctcg ccaggtgtg ggtggttctg gccaggaagg cacaaggtag ctgtgggcca 1500
 agacaccagc cctgtcttag cccttcagta agacttgcg aggaaggagc aaggatgct 1560
 ggggtgccag caagacaagc ccctcagcag gagagaggcc cagaggtccc agctggccac 1620
 cgtgcgccac aagatggccc ctgtggtgtt cctttacg ctggctctct gccagtcctc 1680
 tgcctctcca cctgaacct gcttctctgg ccagtcacag gttggatcc ctctgcatag 1740
 ctgactactc atgcattgct caaagctggc ttctcacatt aagtcaaac caaacctgtg 1800
 tgccacattt catcagagac acacctccct ctggagatgc agttgagtg caaccttgtt 1860
 acattgtatg ctagacaaat tctgtgtgga tatttaagtg aacatgttta caattttgt 1920
 atatatcact cctgacctct cctgaagaag cagagattgt gtattttcag tctcccatgt 1980
 tccagctgca cctctttacc aataaagact gtaactgagc tgactgtgaa aaaaaaaaaa 2040
 aaaaaaaaaa aaaa 2054

<210> 142
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004209

<400> 142
 gatgcagtg agtgacaacc ttgttacatt gtagectaga ccaattctgt gtggatattt 60

<210> 143
 <211> 1224
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004217

<400> 143
 ggccgggaga gtagcagtg cttggacccc agctctcttc cccctttctc tetaaggatg 60
 gcccaagaag agaactccta cccctggccc tacggccgac agacggctcc atctggcctg 120
 agcacccctgc ccagcgagtg cctccggaaa gagcctgtca ccccatctgc acttgctctc 180
 atgagccgctg ccaatgtcca gccacacagt gccctgtggc agaaggtgat ggagaaatgc 240
 agtgggacac ccgacatctt aacgcggcac ttcaacaatt atgactttga gattggggct 300
 cctctgggca aaggcaagtt tggaaacgtg tacttggtct gggagaagaa aagccatttc 360
 atcgtggcgc tcaaggtcct ctccaagctc cagatagaga aggaggcggt ggagcatcag 420
 ctgcgcagag agatcgaaat ccaggcccac ctgcaccatc ccaacatcct cgtctctcac 480
 aactattttt atgacgggag gaggatctac ttgattctag agtatgcccc ccgcggggag 540
 ctctacaagg agctgcagaa gagctgcaca ttgacgagc agcgaacagc cagcatcatg 600
 gaggagtggc cagatgtctct aatgtactgc catgggaaga aggtgatcca cagagacata 660
 aagccagaaa atctgtctct agggctcaag ggagagctga agattgctga ctccggctgg 720
 tctgtgcatg cgccctccct gaggaggaag acaatgtgtg gcaacctgga ctacctgccc 780
 ccagagatga ttgagggcg catgcacaat gagaaggtgg atctgtgtgt cattggagtg 840
 ctttctcatg agctgtcgtt ggggaaccca ccttttgaga gtgcacaca caacgagacc 900
 tatcgccgca tctgtcaaggt ggacctaaag ttccccgctt ctgtgccac gggagccagc 960
 gacctcatct ccaactgctc caggcataac cctcggaac ggctgcccc ggcccaggct 1020
 tcagccaccc cttgggtccg ggccaactct cggagggtgc tgccctctca tgccctcaa 1080
 tctgtcgctc gatggctcct gtcattcact cgggtgcgtg tgtttgtatg tctgtgtatg 1140
 tataggggaa agaagggatc cctaactgtt ccttatctg ttttctacct cctcctttgt 1200
 ttaataaagg ctgaagcttt ttgt 1224

<210> 144
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004217

<400> 144
 gtctgtgtat gtatagggga aagaaggat ccctaactgt tcccttatct gttttctacc 60

<210> 145
 <211> 983
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004335

<400> 145
 gtggaattca tggcatctac ttctgtatgc tattgcagag tgcccatgga agacggggat 60
 aagcgtgtga agcttctgct ggggatagga attctggtgc tctgtatcat cgtgattctg 120
 ggggtgccct tgattatctt caccatcaag gccacacagc aggcctgccg ggacggcctt 180
 cgggcagtgga tggagtgtcg caatgtcacc catctcctgc aacaagagct gaccgaggcc 240
 cagaagggtt ttcaagatgt ggaggccacg gccgccacct gcaaccacac tgtgatggcc 300
 ctaattgctt ccttggtatgc agagaaggcc caaggacaaa agaaagtgga ggagcttagc 360
 ggagagatca ctacattaaa ccataaagctt caggacgcgt ctgcagaggt ggagcgactg 420
 agaagagaaa accaggtctt aagogtgaga atcgcggaca agaagtacta cccagctccc 480
 caggactcca gctccgtcgc ggcgccccc ctgctgattg tctgtctggg cctcagcgct 540
 ctgctgcagt gagatccag gaagctggca catcttgaa ggtccgtctc gctcgctttt 600

```

tcgtctgaac attcccttga tctcatcagt tctgagcggg toatggggca acacgggttag 660
cggggagagc acggggttagc cggagaaggg cctctggagc aggtctggag gggccatggg 720
gcagtccctgt gtgtggggcg acagtcgggt tgaccacagg ctgtctccct ccagagccctc 780
cctccggaca atgagtcctcc cctctgtctc cccaccctga gatggggcat ggggtgcggt 840
gtgggggggca tgtctgtcct gttgttatgg gttttttttg cgggggggggt tgcttttttc 900
tggggtcttt gagctccaaa aaataaacac ttccttttag ggagagcaaa aaaaaaaaaa 960
aaaaaaaaaa aaaaaaaaaa aaa 983

```

```

<210> 146
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004335

```

```

<400> 146
ggttgctttt ttctggggtc tttgagctcc aaaaaataaa cacttctctt gagggagagc 60

```

```

<210> 147
<211> 3446
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004336

```

```

<400> 147
ttctagtttg cgggttcagg ttgccgctgc cggccagcgt cctctggcca tggacacccc 60
ggaaaaatgc cttcagatgc ttgaagccca catgcagagc tacaagggca atgacctctc 120
tggtgaaatg gaaagataca tacagtggtg agaagagaat tttcctgaga ataaagaata 180
cttgataaact ttactagaac atttaatagaa ggaattttta gataagaaga aataccacaa 240
tgaccccaaga ttcatcagtt attgtttaaa atttgtctgag tacaacagtg acctccatca 300
attttttgag tttctgtaca accatgggat tggaaacctg tcatccctc tgatcattgc 360
ctggggcggg catctggaag cccaaggaga gctgcagcat gccagtctgt tctctagag 420
aggaattcaa aaccaggctg aacccagaga gtctctgcaa caacaataca ggttatttca 480
gacacgcctc actgaaaccc atttgccagc tcaagctaga acctcagaac ctctgcataa 540
tgttcagggt ttaaatcaaa tgataacatc aaaaatcaaat ccaggaaata acatggcctg 600
cattttctaag aatcagggtt cagagctttc tggagtgata tcttcaagctt gtgataaaga 660
gtcaaatatg gaacgaagag tgatcacgat ttctaaatca gaatatcttg tgactctac 720
tttggcatcc aaagttgatg ttgagcaggt gtgtatgtat tgcaaggaga agcttattcg 780
tggggaatca gaattttcct ttgaagaatt gagagccagc aaatacaact aacggagaaa 840
gcatgagcaa tgggttaaat aagacagaca ttatatgaaa aggaanaagag caaatgcttt 900
tgaagaacag ctattaaaac agaaaatgga tgaacttcac aagaagtgc atcagtggtg 960
ggagacatcc catgaggatc tgcccgcctc ccaggaaaag tccagaggtt atccagcagc 1020
tatggggcca agtctaggct cccagcagga actgagagcg ccatgtcttc cagttaaccta 1080
tcagcagcaa ccaagtgaaca tggaaaagaa cccaagagag gcacctctgt tgttctctcc 1140
tttggcaaat gctatttctg agacagtaac agactccatg ttgcaatgga ccagcgaaga 1200
tctgttctct ttgaaagccc gtactcatga attcaagcca cagagtgagg cagagatcaa 1320
agaagagtggt gtaacacata aggttgccaa cacaaagtct ttctacacaa ctcacaacac 1380
atcactggga atggttcagg caacgcctac caaagtgcag ccatcaccca cgtgcacac 1440
aaaagaagca ttaggttcca tcatgaatat gtttctaggct cctacaactc ctgtatcttc 1500
tgatgacaaa gatgaatggc aatctctaga tcaaaatgaa gatgcatttg aagccagttt 1560
tcaaaaaaat tgaaggtcat ctggggcctg gggagtcaat atgactatct ctcttttgtc 1620
atctgtcttt catgtgtttg aagatggaaa caaagaaaaa tatggattac cacagcctaa 1680
aaataaaccc acaggagcca ggcacctttg agaacgctct gtacagcagc tctcttcaaa 1740
accaaaggag gaagtgcctc atgctgaaga gtttttggat gactcaactg tatggggtat 1800
tctgtgcaac aaaaacctgg caccagctcc taagagccca ggagacttca catctctgcg 1860
acaacttgcg tctacacatt tccagtggag tccagtgcac tttagaaga 1920
taaagaaaaa gtggtagcaa aacagtgtac ccaggcgact ttgattcttc gtgaggaaaa 1980

```



```

catggtgggtg ccttcaaggg atggaaaatt cagtccaatt caagagaaaa gcccaaaaaa 2040
ggccttgtcg tctcacatgt attcagcatc cttacttcgt ctgagccagc ctgctgcagg 2100
tggggtactt acctgtgagg cagagttggg cgttgaggct tgcagactca cagacactga 2160
cgctgccatt gcagaagatc caccagatgc tattgctggg ctccaagcag aatggatgca 2220
gatgagtcca ctggggactg ttgatgtcc ccaacttcatt gttgggaacc catgggatga 2280
taagctgatt ttcaaacctt tatctgggct ttctaaacca gtgagttcct atccaaatc 2340
ttttgaatgg caatgtaaac ttccagccat caagcccaag actgaatttc aattgggttc 2400
taagctggct tatgtccatc acctctctgg agaaggagcc ttgcccagg tgtacgaagc 2460
taccacaggga gatctgaatg atgctaaaaa taacacagaaa tttgttttaa aggtccaaa 2520
gcctgccaac ccctgggaat tctacattgg gaccagttg atggaagac taaagccatc 2580
tatgcagcac atgtttatga agttctattc tgcccactta ttccagaatg gcagtgattt 2640
agtaggagag ctctacagct atggaacatt attaaatgcc attaacctct ataaaaatc 2700
ccctgaaaaa gtgatgcctc aaggtcttgt catctctttt gctatgagaa tgtctttacat 2760
gattgagcaa gtgcattgact gtgaatcat tcatggagac attaaaccag acaatttcac 2820
acttggaac ggatttttgg aacaggatga tgaagatgat ttatctgctg ccttggcact 2880
gattgacctg ggtcagagta tagatatgaa accttttcca aaaggaaacta tattcacagc 2940
aaagtgtgaa acatctgggt ttcatgtgtg tgagatgctc agcaacaaac catggaacta 3000
ccagatcgat tactttgggg ttgctgcaac agtatattgc atgctctttg gcacttacat 3060
gaagtgtgaa aatgaaggag gagagtgtaa gcttgaaagt ctttttagaa gcttctctca 3120
tttgatgatg tggatgaat ttttctatgt tatgttgaat attccagact gtcatcatct 3180
tccacttttg gatttgttaa ggcaaaagct gaagaaagta ttcaacaac atcatactaa 3240
caagattagg gcctacgta ataggctaat tgtactgctc ttagaatgta agcgttcacg 3300
aaaaaaaatt ttgatatag acagtcctta aaaaacacac tgtaaatag atctgctca 3360
ctttaaacct gtttttttt cattttattg ttatgtaaat gtttgttaaa aataaatccc 3420
atggaatatt tccatgtaaa aaaaaa 3446

```

<210> 148

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004336

<400> 148

ttagggccct acgtaatagg ctaattgtac tgctcttaga atgtaagcgt tcacgaaaaa 60

<210> 149

<211> 739

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004345

<400> 149

```

taaagcaaac ccagccccc accctggcag gcagccaggg atgggtggat cagggaaggct 60
ctgggttggg cttttgcatc aggcctcaggc tgggcataaa ggagctcct gtgggctaga 120
ggagggcaga catggggacc atgaagacc aaagggatgg ccactccctg gggcggtggt 180
cactgtgtct cctgtctgct ggcctgtgta tgccctctggc catcattgcc caggtctoca 240
gctacaagga agctgtgctt cgtgctatag atggcctcaa ccagcggtcc tcggtatgcta 300
acctctaccg cctcctggac ctggaccoca gccccacgat ggatggggag ccagacacgc 360
caaaagctgt gagcttcaca gtgaaggaga cagtgtgccc caggacgcga cagcagtcac 420
cagaggtatg tgacttcaag aaggacgggc tggtagaagg gtgtatgggg acagtgacc 480
tcaacaggcg caggggctcc tttgacatca gttgtgataa ggataacaag agatttgccc 540
tgctgggtga ccaaggatttt ttgcggaaatc ttgtaccagc gacagagtc ttgtgtgtgc 600
tccagagaaat ctacaggtctc tgggctctga gaaataaac atgagagcaa tttcaaaaaa 720
aaaaaaaaa aaaaaaaaa 739

```

<210> 150
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004345

<400> 150
 gcaaagagtt taaaagaatt gtccagagaa tcaaggattt ttgctggaat cttgtaccca 60

<210> 151
 <211> 1432
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004577

<400> 151
 gaggaataatt cttccagcga tggctctcca ctccagagctg aggaagcttt tctactcagc 60
 agatgctgctg tgtttttgatg ttgacagcac ggtcatcaga gaagaaggaa tcgatgagct 120
 agccaaaatc tgtggcggtt aggacgcggt gtccagaaatg acacggcgag ccatgggagg 180
 ggcagtgctct tccaagctg ctctccacaga gcgcttagcc ctcatccagc cctccaggga 240
 ccaggtgcag agactcatag cagagcaacc cccacacctg acccccgga taaggagct 300
 ggttaagtcgc ctacaggagc gaaatgttca ggttttccta atatctgggt gcttttaggag 360
 tattgtagag catgttgctt caaagctcaa tatccacaga accaatgtat ttgccaatag 420
 gctgaataatc taacttaacg gtgaatatgc aggttttcat gagacgcagc caacagctga 480
 atctggtgga aaaggaaaag tgattaaact tttaaaggaa aaatttcatt ttaagaaaat 540
 aatcatgatt ggaagtgggt ccacagatat ggaagcctgt cctcctgctg atgtcttcat 600
 tggatttgga ggaatgtga tcaggcaaca agtcaaggat aacgccaaat ggtatatcac 660
 tgattttgta gagctgctgg gagaactgga agaataacat ccattgtcgt acagctccaa 720
 acaacttcag atgaattttt acaagtata cagattgata ctgtttgctt acagttgcct 780
 attacaactt gctatagaaa gttggtacaa atgatctgta ctttaaaacta cagttaggaa 840
 tcctagaaga ttgctttttt ttttttttta actgtagtct cagtattata tgatgaactat 900
 tgatttcctg gagaggtttt tttttttttt gagacagaat cttgctctgt tgcccaggct 960
 ggagtgcagt ggcgcggtct cggctcaactg caagctctgc ctcccagggt cagcccatc 1020
 tcctgcctca gcctcccgag tagctgggac tacaggcaac cgccaccaca tccggctaact 1080
 tttttgtatt ttttagtagag accgggtttg accgtgttag ccaggatggt ctgtatctcc 1140
 tgacctgtgt atcccctgc ctacagctcc caaagtgcgt ggattacagg ctggggccac 1200
 cgcgcaccag caatgtccta gagagttttg tgatctgaat tctttatgta tattttagc 1260
 tatatttcat acaaaagtgc ttaagtgtgg agagtcaatt aaacacctt acctctagaa 1320
 atacggatct gccagccttc agtgaattat ggtttctctt tggatgtca ataaaagttt 1380
 atccgtatgt cagaacggat ttgtgaaaa aaaaaaaaaa aaaaaaaaaa aa 1432

<210> 152
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004577

<400> 152
 tagaaatcac gattcgycag ccttcagtga atattgggtt ctcttttgga tgtcaataaa 60

<210> 153
 <211> 1530
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_004701

<400> 153

```

aatcctggaa caaggtctaca ggcgtgaaga tccccagcgc tgcgggctcg gagagcagtc 60
ctaacggcgc ctctgtacgct agtgtcctcc cttttcagtc cgcgtccctc cctgggcggg 120
gtctggcactc ttgccttccc cgctccctcat ggcgtctgctc cgacgcccca cggtgtccag 180
tgatttggag aaatttgaca caggagtttaa ttctaaggtt aagagtcatt tgactattag 240
gcgaactggt ttagaagaaa ttggaaatag agttacaacc agagcagcac aagtagctaa 300
gaaagctcag aacaccaaag ttccagttca acccccaaa acaacaaatg tcaacaaaca 360
actgaaacct actgcttctg tcaaaaccagt acagatggaa aagttggctc caaagggtcc 420
ttctccacca cctgaggatg ttcccatgaa ggaagagaat ctctgccaat ctttttctga 480
tgccctgtctc tgcaaaatcg aggacattga taacgaagat tgggagaacc ctccagctctg 540
cagtgactcac gttaaaggata tctatcagta tctcaggcag ctggaggttt tgacgtccat 600
aaaccacatc ttcttagatg gaagagatat aaatggagc atcgctgcc tcttagtgga 660
ttggctggta caagtcact ccaagtttag gcttctgcag gagactctgt acatgtgcgt 720
tggaattatg gatcgatttt tacaggttca gccagtttcc cggaaagaagc ttcaattagt 780
tgggattact gctctgtctc ttggttccaa gtatgaggag atgttttctc caaatattga 840
agactttggt tacatcacag acaatgctta taccagttcc caaatccgag aaatggaaac 900
tctaattttg aaagaattga aatttgagtt gggctgaccc ttgccactac acttcttaag 960
gcgagcatca aagccggggg aggttgatgt tgaacagcac actttagcca agtatttgat 1020
ggagctgact cctctcgact atgatattgt gcattatcat ccttctcaagg tagcagcagc 1080
tgctctctgc ttgtctcaga aggttcttagg acaaggaaaa tggaaactta agcagcagta 1140
ttacacagga tacacagaga atgaagtatt ggaagtcatt cagcacatgg ccaagaatgt 1200
ggtgaagta aatgaagaat taactaaatt cctgcacctc aagaataagt atgcaagcag 1260
caaacctctg aagatcagca tgatccctca gctgaactca aaagccgtca aagaccttgc 1320
ctcccactcg ataggaaggt cctagctctc cgtgggccct ggggatgtgt gcttcatgtg 1380
gccctttttc ttattggttt agaactcttg attttgtaca tagtccctcg gtctatctca 1440
tgaaacctct tctcagaoca gttttctaaa catatttga ggaataata agcgattggt 1500
ttttcttaag tgaaaaaaaaa aaaaaaaaaa 1530

```

<210> 154

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004701

<400> 154

```

agaactcttg attttgtaca tagtctcttg gtatattctca tgaaacctct tctcagacca 60

```

<210> 155

<211> 2536

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004702

<400> 155

```

agcgggtgag gggcggggacc gggccggcct atatattggg ttggcggcgg cgccagctga 60
gccagcgggt agctggtctg gcgaggtttt atacacctga aagaagagaa tgtcaagacg 120
aagtagccgt ttacaagcta agcagcagcc ccagcccagc cagacggaat cccccaaaga 180
agccagata atccaggcca agaagaggaa aactaccagc gatgtcaaaa gaagctcggc 240
taaacattgt aaaaaaggag agcagatatg ttcatgacaa acattttgaa gttctgatt 300
ctgaacttga accacagatg aggtccatc ttctagactg tcttttagag gttatggaag 360
tatacaactc tcatagggaa acatttttat ttgcacaaga cttttttgat agatttatgt 420
tgacacaaaa ggatataaat aaaaatatgc tccaactcat tggaaattac tcattatcca 480
ttgcttccaa acttgaggaa atctatgtct ctaaacctca agagtttgtc tacgtcactg 540
atggtgcttg cagtgaagag gatatcttaa ggaaggaaat cattatatca aaggctttaa 600

```

```

aatgggaact ttgtcctgta acaatcatct cctggctaaa tctctttctc caagttgatg 660
ctcttaaaaga tgcctcctaaa gtctctctac ctacgtattc tcaggaaaca ttcattcaaa 720
tagctcagct tttagatctg tgtattctag ccattgattc attagagttc cagtacagaa 780
tactgactgc tgcgtccttg tgccatttta cctccattga agtgggttaag aaagcctcag 840
gtttggagtg ggacagtatt tcagaatgtg tagattggat ggtacctttt gtcaatgtag 900
taaaaagtac tagtccagtg aagctgaaga cttttaagaa gattcctatg gaagacagac 960
ataatattcca gacacataca aactatttgg ctatgctgga ggaagtaaat tacataaaca 1020
ccttcagaaa agggggacag ttgtcaccag tgtgcaatgg aggcattatg acaccaccga 1080
agagcactga aaaaaccacca ggaaaacact aaagaagata actaagcaaa caagtggaa 1140
ttcaccaga ttgggtagaa ctgggtatcac tgaactacta aagttttaca gaaagttagt 1200
ctgtgattga ttgccctagc caattccaaa gttacactgc cattctgatt ttaaaactta 1260
caattggcac taaagaatcac atttaattat ttctatgtt agctgtlaaa gaaacagcag 1320
gactctgtta caaagatgtc ttcatcccca aggttaactg gttatcttg ataaactagg 1440
ataccoatgca aatgtttttc ctttaatcca gtgttaactg ttatctctg ataaactagg 1500
aahtttgtca ctggagtttt ggactggata agtgcatact taagggtat actaagtgat 1560
acagtacttt gaactcattt gttagattct caaaaactc acactcttga ctagtgcatt 1620
ttggttcttg aaaaataaat taaaacttgt ttacaaaagt ttagttttg aataagtgta 1680
ctaatttctc tatagctgct atagcaagct attataaaac ttgaattct acaaaagtg 1740
aaatttaagt ttttttaaac tagtttattt gccttgcact aacacatttt ttaactaata 1800
aggcttagat gaacatgggt ttcaacctgt gctctaaaca gtgggagtag caaagaaatt 1860
ataaacaaga taaatgctgt ggctccttcc taactggggc ttctctgaca tgtaggttgc 1920
ttggttaataa cctttttgta tatcacaatt tgggtgaaaa acttaagtag cctttcaaac 1980
tatttataty aggaagtcac ttactactc taagatatcc ctaaggaaatt tttttttta 2040
attttatggt actaaggctt tatttatgtt tgtgaaactg ttaaggctct ttctaaattc 2040
ctccattgtg agataaggac agtgtcaagg tgataaagct taacacttga cctaaacttc 2100
tattttctta aggaagaaga gtattaaata tatctgact cctagaatc tatttattaa 2160
aaaaagacat gaaaacttgc tgtacatagg ctactattt ctaaaattt taaatttagt 2220
tttctaaaaa aaaaatccag cctcataaag tagattagaa aactagattg ctagtattat 2280
ttgtatccag atatgtgaat cctctctccc ttggaagaa ctatacatt atgtttacgg 2340
tatgaagtct tctgtatagt ttgtttttaa actaatattt gtctcagtat ttgtctgaa 2400
aagaaaacac cactaatgt gtacatatgt attatataaa cttaacctt taactactgt 2460
tatttttagc ccattgttta aaaaataaaa gttcaaaaaa ttaactgct taaaagtaaa 2520
aaaaaataaa aaaaaa 2536

```

<210> 156

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004702

<400> 156

```

gtttgtgaaa ctggttaagg cctttctaaa ttctccatt gtgagataag gacagtgtca 60

```

<210> 157

<211> 1491

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004710

<400> 157

```

gcggcgccgg cagcgccggc gacggcgaca tggagagcgg ggcctacggc gcggccaagg 60
cgggcgccgt ctcgcagctg cggcgcttcc tgacgcagcc gcaggtgggt gcgcgcgccg 120
tgtcttggt cttgcctctg atcgtgttct cctgcatcta tggtaggggc tacagcaatt 180
cccacagatc taagcagatg tactgcgtgt tcaaccgcga cgaaggatcc tccgcgtatg 240
gcagtgccat cgggtgctg gccttctcgg cctcggcctt ctctctgggt gtcgacgcgt 300
atttccccta gatcagcaac gccactgacc gcaagtaact ggtcatgtg gaactactgt 360
tctcagctct ctggaccttc ctgtggtttg ttggtttctg cttctctacc aaccagtggy 420

```

cagtcaccaa	cccgaaggac	gtgctgggtg	gggccgactc	tgtgagggca	gccatcacct	480
tcagctctct	ttccatcttc	tctctgggtg	tgtctggcctc	cctggccctac	cagcgctaca	540
aggctggcgt	ggacgaactt	atccagaatt	acgttgaccc	cactccggac	cccaacactg	600
cctacgcctc	ctaccacagt	gcattctgtg	acaactacca	acagccaccc	ttcacccaga	660
acgcggagac	caccgagggc	taccagccgc	ccctctgtga	ctgagcggcg	gttagcggtg	720
gaaggggggc	agagagggcc	ctccctctcg	ccctggactt	tcccatgagc	ctcctggaac	780
tgccagcccc	ttcttttcc	ctgttccatc	ctgtgcagct	gacacacagc	taaggagcct	840
catagcctgg	cggggcgtgg	cagagccaca	ccccagtgcc	ctgtgcccag	agggtcttgc	900
tcagccgctc	actctccag	ggcatttcta	ggaaagggtt	ttcagctagt	gtttttcttc	960
gcttttaagt	acctcagccc	cgctgcagct	ggctagaagc	cagcaggtgc	catgtgtcta	1020
ctgacaagtg	cttcagcttc	cccccgccgc	gggtcaggcc	gtgggagccg	ctattatctg	1080
cgttctctgc	caaaagactcg	tgggggccat	cacacctgcc	ctgtgcagcg	gagccggacc	1140
aggctcttgg	gtctcactc	aggtttgctt	ccctctgtgc	cactgtctga	tgtctggggg	1200
gccacacccc	tgtgcgggtg	gcctctgggc	tgccctcccg	ggtgtgaggg	cggggctggt	1260
gctcatggca	cttctctctt	gctccacccc	ctggcagcag	ggaaaggcct	tgccctgacaa	1320
caccacgctt	tatgtaaata	ttctgcagtt	gttacttagg	aagcctgggg	agggcagggg	1380
tgcccatagg	ctccagagct	ctgtctgtgc	cgagtgattg	ataaaatcgt	gggggagatg	1440
cccgccctgg	gatgctgctt	ggagacggaa	taaatgtttt	ctcattcagt	a	1491

<210> 158

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004710

<400> 158

ttgcctgaca acacccagct ttatgtaaat attctgcagt tgttacttag gaagcctggg 60

<210> 159

<211> 3324

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004856

<400> 159

gcagagcacc	gcgccttagc	ccgcaagttc	tagttcttgc	tgccggtcct	aacgtcccgc	60
agtccttcgc	agccagccgt	cccgcatgcg	cgtttggggc	gcgtggagcc	tgctgccatg	120
aagtcagcga	gagctaagac	accccggaac	cctaccgtga	aaaaagggtc	ccaaacgaac	180
cttaaaagacc	cagttggggg	atactgtagg	gtgcgcccac	tggtcttccc	tgtatcaagag	240
tgttgcatag	aagtgatcaa	taatacaact	gttcagcttc	atactctga	gggtcacaga	300
ctcaaacgaa	atggagacta	taaggagact	cagtatccat	ttaaacaagt	tgacctcatt	360
acacccaccc	agaaggaaat	ctttgatgtt	gtggctaate	ccttggtcaa	tgacctcatt	420
catggcaaaa	atggtctctt	ttttcacata	ggtgtgacgg	gaagtggaaa	aactgcacaa	480
atgactgggt	ctccagggga	aggagggtcg	cttcctcggt	gtttggacat	gatctttaac	540
agtatagggt	catttcaagc	taaacgatag	gttttcaaat	ctaatgatag	gaatagtagt	600
gataacaggt	gtgaggttga	tgccctatta	gaacgtcaga	aaagagaagc	tatgcccaat	660
ccaaagactt	cttctagcaa	acgacaagta	gatccagagt	ttgcagatat	gataactgta	720
caagaattct	gcaaagcaga	agaggttgat	gaagatagtg	tctatggtgt	atttgtctct	780
tatattgaaa	tataataata	ttacatafat	gatctattgg	aagaggtgcc	gtttgacccc	840
ataaaaccca	aacctccaca	atctaaattg	cttcgtgaag	ataagaacca	taacatgtat	900
gttcaggagt	gtacagaagt	tgaagtgaac	tctactgagg	aggcttttga	agttttctgg	960
agagggccga	aaaagagacg	tattgtcta	accatttga	atcgtgagtc	cagccgtctc	1020
catagcgtgt	tcacacattaa	attagttcag	gtccctctgg	atgcagatgg	agacaagtgc	1080
ttacaggaaa	aagaacaaat	cactataaagt	cagttgtctc	tggtagatct	tgctggaagt	1140
gaagaaacta	accggaacag	acgacaaagg	aacagattac	gtgaagctgg	taattattat	1200
cagtcactaa	tgacgtctaa	aacatgtatg	gatgtcctaa	gagagaacca	aatgtatgga	1260
actaacaaga	tggttccata	tcgagattca	aagttaaccc	atctgttcaa	gaactactct	1320

gatggggaag	gaaaagtgcg	gatgatcggt	tgtgtgaacc	ccaaggctga	agattatgaa	1380
gaaaactctgc	aagtcatcag	atttgcggaa	gtgactcaag	aagttagaag	agcaagacct	1440
gtagacaagg	caatatgtgg	tttaacgcct	gggaggagat	acagaaacca	gacctgaggt	1500
ccagttggaa	atgaaccatt	ggttactgac	gtggttttgc	agagttttcc	acctttgccg	1560
tcatcgaaa	ttttggatat	caacgatgag	cagacacttc	caaggctgat	tgaagcotta	1620
gagaaaacgac	ataacttacg	acaaatgatg	attgatgagt	ttacaaaca	atctaagtct	1680
tttaaaagctt	tgtttacaaga	atttgacaat	gctgttttaa	gtaaagaaaa	ccacatgcaa	1740
gggaaactcaa	atgaaaagga	gaagatgato	tcaggacaga	aattggaaat	agaacgactg	1800
gaaaaaaaa	acaaaacttt	agaatataag	attgagattt	tagagaaaa	aactactatc	1860
tatgaggaag	ataaacgcga	tttgcaacag	gaacttgaaa	ctcagaaaca	gaactctcag	1920
gcagagtttt	ctgacaaaacg	cagattagaa	gccaggttgc	aaggcatggt	gacagaaacg	1980
acaatgaagt	gggagaaaaga	atgtgagcgt	agagtggcag	ccaaacagct	ggagatgcag	2040
ataaaactct	gggttaaaga	tgaaaagctg	aaacaactga	aggctattgt	tactgaacct	2100
aaaactgaga	agccagagag	accctctcgg	gagcgagatc	gagaaaaagt	tactcaaga	2160
tctgtttctc	catcacctgt	gcctttactc	tttcaacctg	atcagaacgc	accaccaatt	2220
cgctcccgag	acagccgagc	acgctctgca	ggagacagat	gggtagatca	taagcccgcc	2280
tctaacatgc	aaactgaaac	agtcatgcag	ccacatgtcc	ctcatgccat	ccagtatctc	2340
gttgcaaatg	aaaagggcact	agctaaagtgt	gagaagtaca	tgctgaccoca	ccaggaaacta	2400
gcctccgatg	gggagattga	aactaaacta	attaagggtg	atatattataa	aaacaaggggt	2460
ggtggacaat	ctgttcagtt	tactgatatt	gagactttaa	agcaagaatc	accaaatggt	2520
agtcgaaaa	gaagatcttc	cacagtagca	cctgcccaac	cagatggtgc	agagtctgaa	2580
tgagccgatg	tagaaaacaag	gtgttctgtg	gctgtggaga	tgagagcagg	atcccagctg	2640
ggagctggat	atcagcatca	cgcacaaccc	aagcgcaaaa	agccatgaac	tgacagctcc	2700
agactcgaaa	gaacattttc	atttgtgtgg	atgattttctc	gaaagccatg	ccagaagcag	2760
tcttccagtg	catcttgtag	aactccagct	ttgttgaaaa	tcacgcagct	cagctacactc	2820
atacactgac	ccagagcaaaa	gctttcccta	tggttccaaa	gacaactagt	attcaacaaa	2880
ctttgtatag	tatatgtttt	gccatatatta	atattaatag	cagaggaaga	ctcctttttt	2940
catcactgta	tgaatttttt	ataatgtttt	tttaaaatatt	atttcattgta	tactttataaa	3000
ctaattcaca	caagtggttg	tcttagatga	ataagggaaga	ctatatctag	atcatgtctg	3060
attttttatt	gtgactttct	cagocctggt	ctgaattttc	taagggttta	taaacaaatg	3120
ctgctattta	ttagctgcaa	gaatgcactt	tgaactattc	tgacaattca	gactttccaa	3180
ataaagatgt	aaatgactgg	ccaataataa	caattttagg	agggtgtttt	gaattctgta	3240
gttatattatt	cactttctga	catttagata	tgccaaaaga	attaaaaatca	aaagcactaa	3300
gaaataaaaa	aaaaaaaaaa	aaaa	3324			

<210> 160

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004856

<400> 160

caaagcttct	cctattggttc	aaagacaact	agtattcaac	aaaccttgta	tagtgtatgt	60
------------	-------------	------------	------------	------------	------------	----

<210> 161

<211> 1536

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004900

<400> 161

acagagcttc	aaaaaaaaag	cgggacaggg	acaagcggtat	ctaagaggct	gaacatgaat	60
ccaagatca	gaaatccgat	ggagcggatg	tatcgagaca	cattctacga	caactttgaa	120
aacgaaccca	tcctctatgg	tcggagctac	acttggtcgt	gctatgaagt	gaaaaaaag	180
aggggcogct	caaatctcct	ttgggacaca	ggggctcttct	gaggccaggt	gtatttcaag	240
ctcagtgacc	acgcagaaa	gtgcttcctc	tcttggttct	ttggcaacca	gctgcctgct	300
tacaagtggt	tcagatcac	ctgggtttga	tcttgagacc	cctgcccaga	ctgtgtgggg	360

```

aagctggcgc aattcctgtc tgagcaccgc aatgtcaccc tgaccatctc tgcgcgcgcg 420.
ctctactact actggggaag agattaccga agggcgctct gcaggctgag tcaggcagga 480
aatgaaggtc cgatcatgga ctatgaagaa ttgcatact gctgggaaaa ctttgtgtac 540
gcccgcgtga agcaattcat gccctgttac aaattcgatg aaacattcac attcctgcac 600
cgacagctaa aggagattct cagatacctg atggatccag acacattcac ttccaacttt 660
aataatgacc ctttggctct tcagcggcgc cagacactact tgtgctatga ggtggagcgc 720
ctggacaatt gcacctgggt cctgatggac cagcacatgg tgccttatgt caacgaggct 780
aagaattctc tctgtggctt ttacggccgc catgoggagc gctttctatg caacgaggct 840
ccttctttgc agtgggaccc ggcccagatc tacagggtca cttggttcat ctcctggagc 900
cctcgtctct cctgggctgc tgcccgggaa gtgcgtgcgt tccttcagga gaacacacac 960
gtgagactgc gcatcttcgc tgcccgcctc tatgattacg accccctata taaggaggcg 1020
ctgcaaatgc tgcgggatgc tggggcccaa gtctccatca tgacctacga tgagtttgag 1080
tactgctggc acacctttgt gtaccgccag gtagtgcctc tccagccctg gtagtgacta 1140
gaggagacaca gcccaagcctc gagtgggagg ctgcgggcca ttctccagaa tcaggggaaac 1200
gaaggatgag gccctcagtc ctaagggaag cagagaacct ggttgagcac cagaataaaa 1260
gatctctctc caagaaatgc aaacagaccg ttaccacca tctccagctg ctccacagaca 1320
ccagcaaaagc aatgtgtctc tgatcaagta gattttttaa aaatcagagt caattaattt 1380
taattgaaaa tttctcttat gttccaagtg tacaagagta agattatgct caatattccc 1440
agaatagttt tcaatgtatt aatgaagtga ttaattgctt ccattattag actaataaaa 1500
cattaagaat cttccataat tgtttccaca aacact 1536

```

<210> 162

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004900

<400> 162

tgctcacaga caccagcaaa gcaatgtgct cctgatcaag tagatttttt aaaatcaga 60

<210> 163

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004988

<400> 163

```

cgtagaggtc ggccgaagga acctgaccga ggctctgtga ggaggcaagg ttttcagggg 60
acaggccaac ccagagagaca ggattccctg gagggccacag aggagcacca aggagaagat 120
ctgectgtgg gttctcatgt cccagctcct gccacacac ctcctgctg cctgcacgag 180
agtcacatgt tctcttgagc agaggagtct gcaactgcaag cctgagggaag ccttgaggcg 240
ccacaacagag gccctggggc ttgtgtgtgt gcaaggctgcc cctcctcctc cctcctcctc 300
ggtcctgggc accttgagg aggtgccacc tgctgggtca acagatcctc cccagagtcc 360
tcaggagacc tccgcctttc ccaactaccat caacttcaact cgacagaggc aaccagtgga 420
gggttccagc agccctggaag agggaggggc aagcacctct ttgtactctg agtctcttgt 480
ccagagcagta atcaactaaga aggtggctga ttgtggttgt gtcacataaa attacaagca 600
agccagaggag ccaggtcacaa aggcagaaat gctggagagt tgagtccctg cagctggctc ttggcattga 660
ctgttttctc gagatcttgc gcaaaagcctc ctatgtcctt gtcactgcc taggtctctc 720
cgtgaaggaa ccagacccca ccggccactc ctatgatgac ctgctgggtg ataactcagat 780
cctggtcatg attgcaatgg agggcgccca tgctcctgag gaggaaatcg gggaggagct 840
gagtgatgat gagggtgatg atgggaggga gcaacagtgc tatggggagc ccaggaaagt 900
gctcacccaa gatttgggtg aggaaaaata cctggagtag cctgaaacca gatctgtgaa 960
tcccgcacgc tatgagttcc tgtgggttcc aagggccctt gctgaaacca gctatgtgaa 1020
agtccctgag tatgtgatca aggtcagtcg aagaattcgc tttttcttcc catcctctgc 1080
tgaagcagct ttgagagagg aggaagaggg agtctagaca tgagttgcag ccaggggcac 1140
tgggaggggg actggggccag tgcaccttcc agggcgcgct ccagcagctt cccctgcctc 1200

```

```

gtgtgacatg aggcccatc ttcactctga agagagcggt cagtgttctc agtagtaggt 1260
ttctgttcta ttgggtgact tggagattta tctttgttct cttttggaat tggtcaaatg 1320
ttttttttta agggatgggt gaatgaactt cagcatccaa gtttatgaat gacagcagtc 1380
acacagttct gtgtatatag ttttaagggt agagtcttgt gttttatcca gattgggaaa 1440
tccattctat ttgttgaatt gggataataa cagcagtgga ataagtactt agaaatgtga 1500
aaaatgagca gtaaaataga tgagataaaag aactaaagaa attaagagat agtcaattct 1560
tgccattata ctcagttctat tctgtaaaaa ttttaaaagt atagcctaac ctggatttcc 1620
ttgcttctct tgagaatgta agagaaaatta aatctgaata aagaattctt cctgttaaaa 1680
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 1722

```

<210> 164

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004988

<400> 164

```

cagattggga aatccattct attttgtgaa ttgggataat aacagcagtg gaataagtac 60

```

<210> 165

<211> 2334

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004994

<400> 165

```

agacacctct gccctcacca tgagcctctg gcagcccttg gtctctggtgc tccctggtgct 60
ggcgtctgcg ttgtgtcccc ccagacagcg ccagtcacac cttgtgctct tccctggaga 120
cctgagaacc aatctcaccc acagcgagct gccagaggaa tacctgtacc gctatggtta 180
cactcgggtg gtgagatgc gtggagagtc gaaaacctct gggcctgcgc tctgtctctt 240
ccagaagcaa ctgtccctgc ccgagaccgg tgagctggat agccgccacg tgaaggccat 300
gcgaacccca cggctcggggg tccagacctt ggccagattc caaacctttg agggcgacct 360
caagtggcac caccacaaca tcacctattg gatccaaaac tactcggaag acctgcccg 420
ggcggtgatt gacgacgect ttgcccgccg cttcgactg ttgagcgcgg tgacgocgct 480
caccttcaat cgcgtgtaca gccgggacgc agacatcgct atccagtttg gtgtcgcgga 540
gcacggagac gggatccct tcgacgggaa gcacgggctc ctgggacacg cctttccctcc 600
tgcccccggc attcagggaq acgcccaatt ggacgatgac gaggtttgtt gctggggcaa 660
ggcgctctgt gtccaactc ggtttggaaa ccagatggc cgcgctgcc acctccctt 720
catctcgagag ggccgctcct actctgcctg caccaccgac ggtcgctccg accgcttgc 780
ctgtgtcagt accacggcca actacgacac cgacgacggg tttggctctt gccccagcga 840
gagactctac acccgggacg gcaatgctga tgggaaaacc tccagatttc catctcatct 900
ccaagcccaa tctactccg cctgcacac ccagcgtcgc tccgacggct accgctggtg 960
gcaccacacc ccacaactac accgggacaa gctcttcggc ttctgccga cccgagctga 1020
ctcagcgggt atggggggca actcggcggg ggaagtgtgc gtcttcccc tcactttctc 1080
gggtaaaggag tactcgacct gtaccagcga gggccgcgga gatgggcgac tctgtgtgcg 1140
taccacctcg aactttgaca gcgaagaaga gtggggcttc tgcccggacc aaggaatacag 1200
tttgttccct gtggcgcgcg atgagttcgg ccacgcgctg gqcttagatc attcctcagt 1260
gcgggagggc ctcattgtacc ctatgtaccg cttcactgag gggcccccct tgcataaaga 1320
cgacgtgaat ggcactccgc acctctatgg tctcgcctc gaacctgagc cagggcctcc 1380
aaccaccacc acacgcgacg ccacggctcc ccgacgggtc tgccccaccg gacccccacc 1440
tgtccccccc tcagagcgcc ccacagctgg ccccacaggt ccccccacag cgtgccccac 1500
aggtcccccc actgctggcc cttctacgcg cactactgtg cctttgagtc cgttggaaga 1560
tgccctgaac gtgaacactc tcgacgccat cgcggagatt gggaaaccag tgattttgtt 1620
caaagatggg aagtactggc gattctctga gggcaggggg agccggcccg agggcccttt 1680
ccttatcgcc gaagaatggc ccgcgctgcc ccgcaagctg gactcgtgtc ttgagagacc 1740
gctctccaaq aagcttttct tcttctctgg tgggtgtaca tgggcgctgc cccagcgctc 1800
ggtgctgggc ccgagggctc tggacaagct gggcctggga gccgacgtgg cccaggtgac 1860

```



```

ggggggccctc cggagtgcca gggggaagat gctgctgttc agcggggcgc gccctctggag 1920
gttcgacgtg aaaggcgaga tgggtggatcc cgggagcgcc agcggatgtg accgatgtt 1980
ccccgggggtg cctttggaca cgcacgacgt cttccagtac cgagagaaag cctatttctg 2040
ccaggaccgc tttctactggc gcgtgagttc ccggagtggg ttgaaccagg tggaccaagt 2100
gggctacgtg acctatgaca tctcgcagtg ccctgaggac tagggctccc gtccgtcttt 2160
cgagtgcgat gtaaatcccc actgggacca accctgggga aggagccagt ttgccggata 2220
caactgggta tttctgtctg gaggaaaggg aggagtgagg gtgggctggg cccctctctc 2280
tcacctttgt tttttgttgg agtgtttcta ataaacttgg attctctaac cttt 2334

```

<210> 166
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004994

```

<400> 166
ggccctctct tctcaccttt gttttttgtt ggagtgtttc taataaactt ggattctcta 60

```

<210> 167
 <211> 5329
 <212> DNA
 <213> Homo sapiens

<220>
 <221> Modified_base
 <222> 1 ... 5329
 <223> n = a,c,g, or t

<300>
 <308> NM_005063

```

<400> 167
gtgggtgtcgg tgtcggcagc atccccggcg cctgctgcg gtcgcggag cccctcggcct 60
ctgttctcct cccctcccg ccttacctc cagcgggag cgcccgcgcc agtcaactcc 120
tcgcactttg cccctgcttg gcagcggata aaagggggct gaggaataac cggcacagtc 180
caccggttgc cagctctagc ctttaaattc ccggtcggg acctccacgc accggctag 240
cgccgacaaac cagctagcgt gcaaggcgcc gcggtcagc gcgtaccggc gggcttcgaa 300
accgcagtc cccggcgacc ccgaactccg ctccggagcc tcagccccc ttgaaagtgtat 360
ccggcgatcg gagagccaag atgcccggcc acttgcgtca ggacgatatc tctagctcct 420
ataccacac caccacatt acagcgccct cctccagggt cctgcagaat ggaggagata 480
agttggagac gatgcccttc tacttggaa acgacattcg cctgatata aaagatgata 540
tatatgacct caccatacaag gataaggaag gcccaagccc caaggttgaa tatgtctgga 600
gaaacatcat ccttatgtct ctgtacact tgggagccct gtatgggac acccttgatc 660
ctcaactcaa gtctacacac tggcttggg ggttattcta ctattttgtc agtgccctgg 720
tcaggctctt tctgatcatt gccacaacaa tggcattcca gaatgatgtc tatgaatggg 840
cctgggacaa ccgtgccacac cacaagtttt cagaacaaca tgcgtgatctc catattccc 900
gacgtggctt tttctctctc acgtggggtt ggcctgctgt gcgcaaacac ccagctgtca 960
aagagaaggg gagtacgcta gacttgtctg agctagaagc tgagaaactg ttgatgttcc 1020
agaggaggtta ctacaaacct ggcttgcctg tgatgtgctt catcctgcc accgttgc 1080
cctgttattt cttgggtgaa acttttcaaa acagtgtgtt cgttgccact ttcttgcgat 1140
atgctgtggt gcttaatgcc acctggctgg tgaacagtgc tgcccacctc ttccgatata 1200
gtccttatga caagaacatt agccccggg agaatatcct gtttaccatt ggagctgtgg 1260
gtgaggcgctt ccacaactac caccactctt ttccctatga ctactctgcc agtgagtacc 1320
ctgggcacat caacttcacc acattcttca ttgattgcac ggccgcctc ggctggcct 1380
atgaccggaa gaaagtctcc aaggccgcc tcttgccag gattaaaga accggagatg 1440
gaaactcaaa gagtggctga gtttgggtc cctcagggtc ctttttcaaa aaccagccag 1500
gcagaggttt taatgtctgt ttataacta ctgaataatg ctaccagatg cttaaagatg 1560
atgatgttaa cccattccag tacagtattc ttttaaaatt caaaagtatt gaaagccaac 1620

```

aactctgcct	ttatgatgct	aagctgat	tattttctct	cttatccctc	ctctcttota	1680
ggcccaattg	ccctcctttc	actttaatcg	ccctccttcc	cttatttgcc	tcccaggcaa	1740
gcagctgtgc	agctctttgt	cagtgctccag	cttccaaagc	ctagacaacc	ttctctagc	1800
ctaaaacgaa	tggtctttgc	tccagataac	ctctcttccct	tgagctgttg	tgagctttga	1860
atgtagtgcc	tgtagctaga	gataaaacag	aatcttctcg	gtagtcctcc	gttgattatc	1920
ctcagcccg	gcttttgcta	gatggaatgg	aaaagcaact	tcatttgaca	caaagcttct	1980
aaagcnaggt	aaattgtcgg	gggagagagt	tagcatgtat	gaatgtaagg	atgagggaag	2040
cgaaggaacc	ctctcgccag	atcagacata	cagctgccta	octaatgagg	actccaagcc	2100
ccaccacata	gcactgttcc	tttctctcct	ggctcggggt	aaaaagtgcc	tgccgtgttt	2160
ggcaatgcta	attcaatgcc	gcaacatata	gttgaggccg	aggataaagg	aaagacattt	2220
taagttttga	gtaaaagtgg	tctctgtcgg	ggaaggggtt	tcttttcttt	ttttctttaa	2280
taacaagggag	atttcttagt	tcatatatca	agaagctctg	aagttgggtg	tttccagaat	2340
tggttaaaaa	agcagctcat	agaattttga	gtattccatg	agctgctcat	tacagttctt	2400
tcctctttct	ctctctgccat	cttcaggata	ttggtttctc	ccctcatagt	aataagatgg	2460
ctgtggcatt	tccaaacatc	caaaaaaagg	gaaggattta	aggaggtgaa	gtcgggtcaa	2520
aaataaaaaa	tatatatcata	tatacatatg	ttagaacgtt	aaactattag	agtatttccc	2580
ttccaaagag	ggattgtttg	aaaaaactct	gaaggagagg	aggaattagt	tggtgatgca	2640
atttctctct	caactgtcga	catgagatgg	agaggctgag	ggacaggatc	tataggcagc	2700
ttctaaagag	gaacttcaca	taggaagggga	cttgagaaca	cgttcagggg	ttgagaaggt	2760
tactgagtga	cttatttgga	gtcttaataa	actagatatt	aggtccactc	atlaattagt	2820
ttccagttct	ctctgaaagt	agtaaaaaact	agaaggctct	ttctccagct	gttgtgcccc	2880
ttcactcatt	tttttttgag	gagaaggggg	tctctgttaa	catctagctc	aaagtataca	2940
aaactgctgg	ggggcagggt	taggaatctc	ttcactaccc	tgattcttga	ttctctggct	3000
taccctgtat	gtcccttttc	tttgaccaga	tctttctctt	ccctgaacgt	tttcttcttt	3060
ccctgtgacg	gcagccctct	ttgtgtgtat	tcagaggcag	tgatgaactg	ctgtccagcg	3120
agctccctcc	tcgacacaga	atgctcaggg	tcactgaacc	actgtctctc	ttttgaaagt	3180
agagctagct	gccactttca	ctgggctctc	gcaggtgtct	ccactacacc	ccctgtctcc	3240
actgcccacc	tgatgggtca	agacaaggct	ggcaaacctc	cccagaaca	ttctctggcc	3300
agaaagcgtc	ttctctccctc	ctctctcat	gagaaagcca	ggcctcatgt	tgagccagtg	3360
ggccagccac	agagcaaaaag	agggtttatt	ttcagtcctc	ttctctctgg	tcagacaacg	3420
agggcatgct	gaatgcccc	tgcttacttg	gtgaggggtc	cccgcctgag	tcagtgctct	3480
cagctggcag	tgcgaatgct	gtagaagtag	gaggaacaag	ttctcaactg	gaaggaagca	3540
gggcaagAAC	ccaagtgcct	caacctgaaa	ggaggccctg	ttccctggag	tcagggtgaa	3600
ctgcaaaagt	ttggctgaga	cctgggattt	gagataccac	aaacctgctg	gaacacagtg	3660
tctgttcagc	aaactaacca	gcattcccta	cagcttaggg	cagacaatag	tatagaagtc	3720
tggaaaaaaa	caaaaaacaga	atttgagaac	ctctgaccac	tctgttccct	tgagctcagt	3780
catcaaaagc	gaagctctggc	tttgctctat	taagatttga	aatgtacact	accaaaacct	3840
cagtcacagt	ttgagcccca	gtgctggaag	ggaggaaagg	ctttctctctg	gtttaattgc	3900
gtagaggcta	cagggtgttag	octggactaa	agggatcctt	gtctttgagc	tattcacctc	3960
agtagaaaaa	gatctaaggag	aagatcactg	tagtttagtt	ctgttgacct	gtgcacctac	4020
cccttgaaaa	tgctctgtgg	tatttctaata	tccacaggtc	atcagatgcc	tgcttgataa	4080
tatataaaca	ataaaaaaaa	ctttcacttc	ttctctattg	aactcgtgtc	catggatctg	4140
attctgtacca	tgacctacaa	taaggctgga	tggaacctca	ggctgagggc	cccaatgtat	4200
gtgtggctgt	gggtgtgggt	gggagtggt	ctgctgagta	aggaacacga	tttccaagat	4260
tctaaagctc	aatctcaagt	acacattaat	gataaaactca	gatctgtaca	agagctccga	4320
tttctaacag	ctctgtgctt	gggggggtgt	ctggcaactc	agctcagggt	ctttacatct	4380
tttctaataca	cagtggtgtca	tatgagcctg	ccctcactcc	ctctgcagaa	cttctttgca	4440
octgagacc	tactgaagtgt	gctgttagaa	aaaggggctc	gagtgaggga	ttatcagtat	4500
caogatttgc	aggatttccct	ctcgggcttc	attctggaaa	cttttgttag	gttctctgtt	4560
cttaagtgcc	cacatttgat	ggaggggtga	aataatttga	atgtatttga	tttataagtt	4620
tttttttttt	tttgggttaa	aagatgggtg	tagcatttga	aattggaatt	tttctctctg	4680
gtttgtcagt	atctctgggtg	tattctctgt	aagtgtagct	caaataggtc	atcatgaagag	4740
gttaaaaaag	cgaagtgccc	atgttatgct	gggtgttgct	agggctccca	accactgtgc	4800
cactgacttg	ctgtgtgacc	ctgggcaagt	cacttaacta	taaggtgctc	cagttttctc	4860
ctctgttaaa	tgggataaat	aatactgacc	tacctcaaag	ggcagttttg	agcgaagact	4920
aatgtctttt	agaaagcatt	ttgggatcct	tcagcacagg	aattctcaag	acctgagtat	4980
tttttataat	aggaatgtcc	accatgaact	tgatacgtcc	gtgtgtccca	gatgctgtca	5040
ttagtctata	tggtttctcca	agaaactgaa	tgaaatccatt	ggagaagcgg	tggaataact	5100
cccagacaaa	atttgagaat	acataaaca	cgcattgtcca	cggaaacata	cagaggaatgc	5160
gttctctgtg	atttgagggt	attttttccc	tttttatgtc	ggatatagta	gttactgtg	5220

acaagaataa ttttgaata atttctatta atatcaactc tgaagctaat tgtactaatc 5280

tgagattgtg ttgttcata ataaaagtga agtgaatctg attgcactg 5329

<210> 168

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005063

<400> 168

aataatgcta ccaggatgct aaagatgatg atgttaaccc attccagtac agtattcttt 60

<210> 169

<211> 634

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005101

<400> 169

cggctgagag gcagcgaact catctttgcc agtacaggag cttgtgccgt ggcccacagc 60
 ccacagccca cagccatggg ctgggacctg acggtgaaga tgetggcggt caacgaattc 120
 caggtgtccc tgagcagctc catgtcgggt tcagagctga aggcgcagat caccacagaag 180
 attggcgtgc acgccttcca gcagcgtctg gctgtccacc cgacgcggtgt ggcgctgcag 240
 gacagggtcc cccctgccag ccagggcctg ggccctggca gcacggctct gctgtgggtg 300
 gacaatgcg acgaacctct gacatcctgt gtgaggaata acaagggccg cagcagcacc 360
 tacgaggtcc ggcctgacga gacgtggccc cactgaagc agcaagtga cgggctggag 420
 ggtgtcagc agcactgttt ctggctgacc ttcgagggga agccctgga ggaccagtc 480
 ccgctggggg agtacggcct caagccctgt agcaccgtgt tcatgaattc gcgcctgcgg 540
 ggaggcgcca cagagcctgg cgggcggagc taagggcctc caccagcacc cgagcaggat 600
 caaggccggg aataaaaggc tgttgaaga gaat 634

<210> 170

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005101

<400> 170

tggtgtgga caaatgcgac gaacctctga gcactcgtgt gaggaataac aagggccgca 60

<210> 171

<211> 1339

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005139

<400> 171

gaattccgat tagtgtgac tcagctcaag gcaagggtgg gatatcatgg catctatctg 60
 ggttggacac cgaggaaacag taagagatta tccagacttt agcccatcag tggatgctga 120
 agctattcac aaagcaatca gaggaattgg aactgatgag aaaatgctca tcagacttct 180
 gactgagagg tcaaatgcac agcggcagct gattgttaag gaatatcaag cagcatatgg 240
 aaaggagctg aaagatgact tgaagggtga tctctctggc cactttgagc atctcatggt 300

```

ggccctagtg actccaccag cagtctttga tgcaaaagcag ctaaagaagt ccatgaagg 360
cgccgggaaca aacgaagatg ccttgattga aatcttaact accaggacaa gcaggcaaat 420
gaaggatata tctcaagcct attatacagt atacaagaag agtcttgagg atgcatttag 480
ttccgaaaca tctgtgtgact tccggaaagc tctgttgact ttggcagatg gcagaagaga 540
tgaaagtctg aaagtggatg agcatctggc caaacaagat gccagatctc tctataaagc 600
tgggtgagaac agatggggca cggatgaaga caaatteact gagatcctgt gttaaaggag 660
ctttctccaa ttaaaactaa catttgatga atacagaat atcagccaaa aggacattgt 720
ggacagcata aaaggagaat tatctgggca ttttgaagac ttactgttgg ccatagttaa 780
ttgtgtgagg aacacgcogc ccttttttag cgaaagact catcgagcct tgaagggtat 840
tggactgat gagtttactc tgaaccgaat aatggtgtcc agatcagaaa ttgacctttt 900
ggacattcga acagagtcca agaagcatta tggctattcc ctatatccag caatataatc 960
ggatacttct ggagctatg aaatcacact cttaaaaatc tgtggtggag atgactgaac 1020
caagaagata atctccaaag gtccacgatg ggcttcccca acagctccac ctactctctt 1080
ctcatctcat taaagagAAC aagcaaatat aaacagcaac ttgtgttctc aacaggaatt 1140
ttcattgttc tataacaaca acaacaaaag cgattattat tttagagcat ctcatattat 1200
atgtagcagc tcataaatga aattgaaat ggtatataag atctcgcaat actatccaac 1260
ttatatctct gcttccaaag ttaagaatct ttatagtctt actccattaa atataaagca 1320
agataataaa acggaattc 1339

```

<210> 172

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005139

<400> 172

ttcagcaatt aaatcggata cttctggaga ctatgaaatc acactcttaa aaatctgtgg 60

<210> 173

<211> 1582

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005165

<400> 173

```

ccgagctgtg cttgtggctg cggctgctaa ctggctgcgc acagggagct gtcaccatgc 60
ctcaectcgt cccagccctt tctgctgagc agaagaagga gttgtctgac attgcccctgc 120
ggattgttagc cccgggcaaa ggcattctgg ctgcggatga tctgttaggc agcatggcca 180
agcggctgag ccaaatgtgg gtgaaacaca cagaggagaa ccgcccgtcg taaccgccagg 240
tccattcagc tgcgtgtgac cgtgtgaaaa agtgcatctg aggcgtcatt ttcttccatg 300
agaccctcta ccagaaagat gataatgggt ttcccttcgt ccgaaccatc caggataagg 360
gcacgtcgtg gggcatcaag gttgacaagg gtgtggtgoc tctagctggg actgatggag 420
aaacaccacc tcaagggtcg gatgggctct tgcccaatag aagaaggatg 480
gtgctgactt tgcgaagtgg cgtgtgtgic tgaaaaatcag tgagcgtaca cccctgtcac 540
ttgcatctct ggagaacgcc aacgtgctgg ccogttatgc cagcagaattg 600
gcatctgtgc tatgtgtgaa cctgaaatat tgcctgatgg agaccacgac ctcaaacgtt 660
tcaglatatg tacagagaag gtcttggctg ctgtgtacaa ggcctgagt gaccatcatg 720
tataacttga ggggaccctg ctcaagccca acatggtgac cccgggccat cgcctgtcca 780
tcaagtatac cccagaggag attgccatgg caactgtcac tgccctgcgt cgcactgtgc 840
ccccagctgt cccaggatgt acccttctgt ctgggggtca gagcgaagaa gaggcatcat 900
tcaactcaca tgcactcaac cgtgcctccc ttcccccacc ctgggcctct acccttctct 960
atgggcgtgc cctgcaagcc tctgcaactc atgctctggc agggcaacgg gacaatgctg 1020
ggcgtgcctc tgaggagttc atcaagcggg ctgaggtgac tgggcttgca gccagggca 1080
agtatgaagg cagtggagaa gatggtggag cagcagcaca ctgcaactac attgccaacc 1140
atgctcactg agtatccact ccataccaca gcccttggcc cagccatctac caccacttct 1200
tgcctgtagt cactggccagg gccaaatagc tatgcagagc agagatgcct tcaactgtga 1260
ccaactgttc ttctcttctc tcttcccttc cctctctcct ttgctgcacc tgggaccata 1320

```

```

ggatggggag ataggagacc cctcatgact gagggcagaa gaaattgcta gaagtcagaa 1380
caggatggct ggggtccccc ctacctcttc cagctccacc aattttccca tgatgaggtta 1440
gctttccctt ggggtctctt tottgctctg cctgtctcct gggatcagag ggtagtacag 1500
aagccctgac tcatgcttg agtacatacc atacagcaaa taaatggtag caaaacaaaa 1560
aaaaaaaaa aaaaaaaaaa aa 1582

```

<210> 174

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005165

<400> 174

```

gagggtagta cagaagccct gactcatgcc ttgagtacat accatacagc aaataaatgg 60

```

<210> 175

<211> 451

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005213

<400> 175

```

acttccctgt tcaatttgggt tocagcatcc tgtccagcaa agaagcaatc agccaaaaatg 60
atacctggag gcttatctga ggccaaaccc gccactccag aaatccagga gattgttgat 120
aaggttaaac cacagcttga agaaaaaaca aatgagactt atggaaaatt ggaagctgtg 180
cagtataaaa ctcaagttgt tgcctggaaca aattactaca ttaaggtacg agcaggtgat 240
aataaatata tgcacttgaa agtattcaaa agtcttcccg gacaaaatga ggaacttgta 300
cttactggat accaggttga caaaaacaag gatgacgagc tgacgggctt tttagcagcat 360
gtacccaagc tgtttcattt ccttcaactg gctactgagt catgatcctt gctgataaat 420
ataaccoatca ataaagaagc attcttttcc a 451

```

<210> 176

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005213

<400> 176

```

aactggctac tgagtcatga tctctgctga taaataaac catcaataaa gaagcattct 60

```

<210> 177

<211> 366

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005218

<400> 177

```

gtcagctcag cctccaaaag agccagcctc tccccagttc ctgaaatcct gagtgttgcc 60
tgccagtcgc catgagaact tctaccttc tgcgtgttac tctctgotta cttttgtctg 120
agatggcttc agatggtaac ttctccacag gccctggcca cagatctgat cattacaatt 180
gcgtcagcag tggaggggcaa tgcctctatt ctgctcgccc gatctttacc aaaattcaag 240
gcacctgtta cagagggaag gccaaagtgc gcaagtgcgc tgggagtgac cagaagaatt 300
gacgcagaag tgaaatgaac tttttataag cattctttta ataaaggaaa attgcttttg 360

```

aagtat 366

<210> 178
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005218

<400> 178
 gggagtgacc agaagaaatg acgcagaagt gaaatgaact ttttataagc attcttttaa 60

<210> 179
 <211> 1519
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005326

<400> 179
 ctgcctcgga acgctgtccc cgcagcgcac ggcccggttc acctcgcgat ctgcgggta 60
 ccggggcgcc gtggcgctcg gcctccaggg atccactgtg cgggtgccaa aaagaggcgg 120
 aggtctcgccg cacagctctc cgggcgcagc tctcgggccc ccgccgccgc tccaggcccc 180
 gtctcccgcc ccgtggcagt cggggctcgc ggacaaaaca agttgagcgc gagcgcgttg 240
 attgtttggc ggacggtgct aggtggacgc tgattggctg agggcacgcg gagcgggcgc 300
 ctgattggct cgcagcgccc gaocggcggt ttttgacgtc ctgggcagct cggcagttcca 360
 gcccgggccc ggtcatggtg gtgggcgcag ggctgctcgg ccgccgcagc ctgcgcgcgc 420
 tgggagcgcg ctgcgccccc cgagggcctcg gtccagccct gctgggagtt ttctgccaca 480
 cagatttggt gaagaacctg accgtggacg agggcaccat gaaggtagag gtgctgcctg 540
 ccttgaccga caactacatg tacctggcca ttgatgatga gaccaaaggag gctgccattg 600
 tggatccggt gcagccccc aggttcgttg acgcggcgag aaagcagcgg gtgaaactga 660
 ccacagtgct caccacccac caccactggg accatgctgg' cgggaatgag aaactgggtca 720
 agctggagtc gggactgaag gtgtacgggg gtgacgacgc tatcgggccc ctgactcaca 780
 agatcactca cctgtccaca ctgcaggttg ggtctctgaa cgtcaagtgc ctggcgaccc 840
 cgtgccacac ttcaggcacac atttgttact tctgtgagca gcccgagggc tgggagcccc 900
 ctgcgctggt cacaggtgac acctgtgttg tggctggctg cgggaagtgc tatgaaggga 960
 ctgcggatga gatgtgtaaa gctctgcttg aggtcttggt ccggtctccc ccggcacaaa 1020
 gagtctactg tggccacgag tacaccatca acaacctcaa gtttgacgc cactgtggagc 1080
 ccggcgaatg cgccatccgg gagaagctcg cctggggcaa ggagaagtag agcatcgggg 1140
 agcccacagt gccatcaccc ctggcagagg agtttaccta caaccccttc atgagagtga 1200
 gggagaagac ggtgcagcag cagcgaggtg agacggaccc ggtgaccacc atgcggggccg 1260
 tgcgcagggga gaaggaccag ttcaagatgc ccgcggactg aggcgcctct gcaacctcag 1320
 cggattttgg gattagcttc ttttaggtaa ctggctttcc tgcgtgtccg tgcgggaaat 1380
 tcatcttga tttaacctta attttacagc ccttggttg tgttatcgga catctcaag 1440
 catatttata agagaagttt aacaagtatt tattccata aaaaaaaaaa 1500
 aaaaaaaaaa aaaaaaaaaa 1519

<210> 180
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005326

<400> 180
 cttgtgttat cggacattct aatgcatatt tataagagaa gtttaacagg tatattatcc 60

<210> 181

<211> 3378
<212> DNA
<213> Homo sapiens

<300>
<308> NM_005461

<400> 181
acagctgcac cgccggagctg cgagcggctg cgagcgagag agcgtaaag caagagagct 60
agagcgcag caacggggcac cctccctca gccccaccgc gcgctccgct 120
tgccctccca ccccgcccca ctctaccggt cccggtccct gcgcgggcac agccagagac 180
ctctggggcgg tgacggcagc ctcgggactc tccggcgccg cgccgctcgc ccagacaaa 240
cgcttggcgg cgcccgccgg ccgctgcgcc ctgcctcccc gccctcccag ctcttctccg 300
ctcttcccc ccgctgttgg ctcgccgcgc tccggccggc cgcaaaagt tt ccggcgccg 360
agcgcggtct gcgcctgctc tcagcgatgg ccgcgagct gagcatgggg ccagagctgc 420
ccaccagccc ctgggccatg gagtatgtca acgacttca cctgctcaag ttgcagctga 480
agaaggagcc actggggcgc gcggagcgctc cgggcaggcc ctgcacacgc ctgcagccag 540
ccggtctcgt gtccctccca ccgctcagca ctccgtgtag ctccgtgcc ctgctgccca 600
gcttcagccc gacccaacag aagacacacc tcgaggatct gtactggatg gcgagcaact 660
accagcagat gaaccccgag gcgctcaacc tgacgccga ggacgcggtg gaagcgctca 720
tcgctgcga cccagtgcca cagccgtctc aaagcttca cagctttcgc ggcgctcacc 780
accaccacca tcaccaccac cctcacccgc accacgcgt cccggggccc ggctgtggccc 840
acgacagctt gggcccgccac gctcacccgc accatccca tcatcccaa gcgtgcgcgc 900
cgccgtccag cgcgcctagc ccggcgcaac agctgccacc tagccacccc gggcccgggc 960
cgcacgcgac ggcctcgccg acggcgccgg gcggcaacgg cagcgtggag gaccgcttct 1020
ccgacgacca gctcgtgtcc atgtccgtgc gcgagctgaa ccgccaacct cgggggttca 1080
ccaagacga ggtgatccgc ctgaagcaga acggcggaac cctgaagaac cggggctacg 1140
cccagctctt caggtataaa ccgctccagc aagaagcaca cctggagaat gagaagacgc 1200
agctcaattc caggtggag cagcttaagc aggaagtgct ccggctggcc cgcgagagag 1260
acgctacaaa ggtcaagtc gcgaactcgc ccaactccgg ctccagggag gcgggctcca 1320
ccagcgacag cccctccctc ccagagttct ttctgtgagt cgtggccggt cctggccccc 1380
gccccttgcc ccggcccgac agctcctag tcccagacta ccccggaacc 1440
tgtccctgcc gcggcccccag ccttgacctg ttgacttga gcgagagga ggaaggcgcc 1500
gcggcccgcc ggcgcagggc ggggtgcgcg gcgggcaggg gaccttgct aaggcgagag 1560
tagcgcaagg cagcgccccc tcttagactc gagcagagcc ggagagagag acgagaggg 1620
gggaggtccc ggaatcaact ctctccagc tgaaggcgcg cgagcgatag tcccgaag 1680
tcaccaaggc catctggaga ctctcggtt tctgaacttt gcgcgttaag ccgggacagg 1740
tgctttgctg cccgcagagt agtcgcgcgc aggaagagag caacgaggaa aggagaggga 1800
ctctgcgctc ccggcagggc agagcgagag ctgagcgaaa gaaggagga cagacggacc 1860
tgtctctcag agtttcgaga acactggctc tcaagccctga gacacaggcc tcaagttaga 1920
cgctcggcgc caaaatctca tcagttttat tgcctgctcg attataaga aaaaatacaa 1980
aaatctgcac taaaataatt aatcctgcac gctggacat tatgtaata atttctatt 2040
tgtacattt tcttgtttaa ctttagcatg ttgttgatca tggatcatac tcccttgtt 2100
tctttgggtt agaaggatc gcagtttga aactccggcg gctgcgtgcg gggtttcagt 2160
cccagctgta ggtctgtaaa tcaccgcccc gccaaaacgc atagagaacg tggcagaggg 2220
ctgagggctc ttggtttgggt ttattattac ggttatgtt ttgtgaagtt aaaaagaaa 2280
aaaaaaagaa aaagtccogg gcaattccgg tcaaaaaa atcttgcct tcttgcctc 2340
tggaagattt atgttttctt ttggtgtgag agagagaacc gactcctct ttctctctc 2400
gctttagttt tcaacactgt tgggtgctgag agagagaacc agtacaaggg 2460
caggcgaggc cagggaagct gccaaagctc gcaacccaga ggagtgtctt ggactcacg 2520
ctgtcttat ataccttaa taagaaggga aaggaaagga aacacagatc 2580
tcccctctgc tttttattgt aaccagaatc accctgaggt cccctctgaa cccctctggc 2640
ctgcgtcaat tgtatggacc acagcgctcc tagggttga gcttatgcca tccctgacc 2700
tggcagtgca ctgtgaacca gacactgcac tgaaccaa gctatgtcca gaattgacca 2760
gaaccccaaa catgtgcaag taattttgca actttcaag cgttcttta gcccaatgca 2820
ttcgctttct ttccctgctt ttgagatagt aggaagagtt ctgtgtggtg tccccccct 2880
tcaatttctc agttgtatag taattatagg gaagatatg gtgtttttct ttattattac 2940
ttttttttt ctgcaggtca gtaaaaggat ttaagttgca ctgacaaaa taccaaaaa 3000
aaagtgtatt ttaagttcc cattgaaat tgctggcggt gctggccgga tgcattttt 3060
agtttgtatt agttgataa ttaacagtaa gatatcaacg catgggtgct 3120
gcagttttta atatttgtga tatttgtcct gcaatcagaa cgagcttttg tttttacaga 3180

```

ttcaactgtg ttgaaatcaa acctgccgca acagaaattg tttttatttc atgtaaaata 3240
agggatcaat ttcaaacccct gcttatgata tgaaaattat aaaacctagt ctattgtagt 3300
tttattcaga ctggtttctg ttttttggtt attaaaaagg tttccatttt tgcttattaa 3360
aaaaaaaaaa 3378

```

```

<210> 182
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005461

```

```

<400> 182
atttgtctcg catcagaaac gagctttggt ttttacagat tcaactgtgt tgaaatcaaa 60

```

```

<210> 183
<211> 597
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005532

```

```

<400> 183
agctgaagtt gaggatctct tactctctaa gccacggaat taacccgagc aggcattggag 60
gcctctgctc tcacctcatc agcagtgacc agtggtgcc aagtggtcag ggtggcctct 120
ggctctgccg tagttttgcc cctggccagg attgctacag ttgtgattgg aggaattgtg 180
gccatggcgg ctgtgcccat ggtgctcagt gccatgggct tcactgcggc gggaatcgcc 240
tcgtctctca tagcagccaa gatgatgtcc gcggcgcca ttgccaatgg ggggtggagt 300
gcctcgggca gccttgtggg tactctgcag tcactgggag caactggact ctccggattg 360
accaagtcca tctctgggtc cattgggtct gccattgcgg ctgtcattgc aggtttctac 420
tagtcctcctg cccctcgccc tgcagagaag agaaccatgc cagggggagaa ggcaccagc 480
catcctgacc cagcgaggag ccaactatcc caaatatacc tgggtgaaat ataccaaatt 540
ctgcatctcc agaggaaaat aagaataaaa gatgaattgt tgcaactctt aaaaaa 597

```

```

<210> 184
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005532

```

```

<400> 184
agccaactat ccCaatatata cctgggtgaa atataccaaa ttctgcattc ccagaggaaa 60

```

```

<210> 185
<211> 1661
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005566

```

```

<400> 185
tgctgcagcc gctgcgcgcg attcoggatc tcattgccac gcgcgccgca cgaccgccg 60
acgtgcattc ccgattcctt ttggttccaa gtccaatatg gcaactctaa aggcacagct 120
gattataaat cttctaaaag aagaacagac ccccagaatg aagattacag ttgttgggtg 180
tgggtgctgtt ggcattgcct gtgcacacag tatcttaatg aaggacttgg cagatgaact 240
tgctcttgtt gatgtcatcg aagacaaatt gaaggggagag atgatggatg tccacatgag 300

```



```

cagccctttc cttagaacac caaagattgt ctctggcaaa gactataatg taactgcaaa 360
ctccaagctg cgtcatatca cggctggggc acgtcagcaa gaggagaaaa gccgtcttaa 420
tttggctcag cgtaacgtga acatatttaa attcatatt cctaattgtt taaaaacacg 480
cccgaactgc aagttgttta ttgtttcaaa tccagtgatg atcttgacct acgtggcttg 540
gaagataagt ggttttccca aaaaccgtgt tattggaagt ggttgcaatc tggattccagc 600
cggatttcgt taccctgatgg gggaaaggct gggagttcac ccattaagct gtcattgggtg 660
tgtctctctg aagactctgc acccagattt agggactgat aaagataaag aacagtggaa 720
agaggttcac aagcaggtgg ttgagagtgc ttatgaggtg atcaaaactca aaggctacac 840
ccttgggctt attgactctc ctgtagcaga ttggcagag agtataatga agaattcttag 900
cggggtgcac ccagtttcca ccatgattaa gggctcttac ggaataaagg atgatgtctt 960
ccttagtggt ccttgcatct ttggacagaa tggaaatctca gacctgtga aggtgactct 1020
gactctcgag gaagaggccc gtttgagaa gagtgcagat acactttggg ggatccaaaa 1080
ggagctgcga ttttaaagtc ttctgatgtc atatcatttc actgtctagc ctacaacagg 1140
attctaggtg gaggttgtgc atgttgtctt ttttatctga tctgtgatta aagcagtaat 1200
attttaagat ggaactggaa aaacatcaac tctgaagtgt agaataaaga atggttttga 1260
aaatccacag ctatatctctg atgctgggat gtattaatct tgtgtagtct tcaactgggt 1320
agtggtgaat agttctgccca cctctgacgc accactgcca atgctgtacg tactgcaatt 1380
gccctctcag gtttacgggt tttaacgggt ttttatataa ctctctggct ccttcactga 1440
acatgctcat tccaacattt tttccagtg agtcacatcc tgggattccag tgtataaatt 1500
caatatcatg tcttgtgcat aattcttcca aaggatctta ttttgtgaac tatatcagta 1560
gtgtacattc ccatataatg taaaagatc tacatacaaa caatgcaacc aactatccaa 1620
gtgttatacc aactaaaacc cccaataaac cttgaacagt g 1661

```

<210> 186
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_005566

<400> 186
catcaactcc tgaagttaga aataagaatg gtttgtaaaa tccacagcta tatctgatg 60

<210> 187
<211> 2993
<212> DNA
<213> Homo sapiens

<300>
<308> NM_005689

```

<400> 187
gggcctgcag ttggcagaag ggtcccgggc ccagagccag cggggccgtg ctgagacggc 60
gtacgtgcc cgcgtgaagt cgtggcgggc ggcgtgcgc taggggagtg ggcggtgagg 120
cctggtccac gtgcgtccct tcccgggacc ccgcagctt ggcgccacg ggctacgtga 180
cgcaaggcac ccgagatgtc cgcgccctct ccgagtgaca agtcccggcc tccggtccag 240
cagtgccgc cagctcgccc ggcgtccacg ggcgtccacg cacttgctgg gtgactgtgg gcaactactg 300
cgagggcgaa gggcccgctg gtcggccctg gatgcaggat ggctgagtc cctgctctct 360
cttcaagctc gtgccctcga cgcggatggc tctaggact ctggccttgg tgcgtgctct 420
tccctcaga cgcggggagc ggcccgctgg tctgatttcg ctgtcttggg gggccggccc 480
tcgactctct cctaactgct tgcagctgct tctggccaca cttcaggcgg cgtcgccctt 540
ggccgctcgt gctggccggg tgggcaactgc ccggggggcc ccaactgccaa gctatctact 600
tctggctccc gctccttgga gtctggccgg cgcctgtggc ctgtggctgc ttgtcgttga 660
cgggacccag gcaaggcagc gtctggcaat gggcactctg atcaagttca ggcacagccc 720
tggtctctg ctcctctgga ctgtggcgtt tgcagctgag aacttgcccc ttgtgtcttg 780
gaacagccca cagtgggtgtt gggcaagggc agaactgggc caacaggttc agtttagcct 840
tggtgtgctg cggatagtgg tctctggagg gctgtttgtc ctgggtctct gggccctctg 900
acttcgtccc cagtctcata cattgcaggt tcatgaagag gaccaagatg tggaaaggag 960
ccaggttcgt tcagcagccc aacagctcac ctggcgagat tttggcagga agctccgctt 1020

```

```

cctgagtggc tacctgtggc ctccagggag tccagctctg cagctgggtg tgctcatctg 1080
cctggggctc atgggtgttg aacgggcact caatgtgttg gtgcctatat tctataggaa 1140
cattgtgaac ttgtgtgact ttggaacctc ttggcctgga ctgtttaccag 1200
ttacgtcttc ctcaagttcc tccagggggg tggcactggc agtacaggct tctgtagcaa 1260
cctgcgcacc ttccctgtga tccgggtgca gcagttcacg tctcgcgagg tggagctgct 1320
catctctctc cacctgcacg agctctcact gcgctggcac ctggggcgcc gcacagggga 1380
gggtgctgcg atcgcggatc ggggcacatc cagtgctcac gggctgtcca gctacctggt 1440
gttcaatgtc atccccagc tggccgacat catcatggc atcatctact tcagcatggt 1500
cttcaacgcc tggtttggcc tcatgtgttt cctgtgcatg agtctttacc tcaacctgac 1560
cattgtggtc actgagtggg gaaccaagtt tcgtcgtgct atgaacacac aggagaaacg 1620
taccggggca cgagcagtggt actctctgct aaacttcgag acggtgaagt attacaacgc 1680
cgagagttac gaagtggaa cgtatcgaga gccatcatc aaatatcagg gtttggagtg 1740
gaagtgcagg gcttcactgg ttttactaaa tcagaccag aaactgggtg ttgggctcgg 1800
gctcctcgcc ggtccctcgc tttgcgcata ctttgcact gaggcagaag tacaagtttg 1860
ggactatgtg ctctttggca cctacattat ccagctgtac atgcccctca attggtttgg 1920
cacctactac aggatgatcc agaccaactt cattgacatg gagaacatgt ttgacttget 1980
gaaagaggag acagaagtga aggaccttcc tggagcaggg ccccttcgct ttcagaaagg 2040
ccgtattgag tttagaagcg tgcacttcag ctatgccgat gggcgggaga ctctgcagga 2100
cgtgtctttc actgtgatgc ctggacagac acttgccctg gtgggcccct ctggggcagg 2160
gaagacacaa attttgcgcc tgcgtgtttc cttctacgac atcagctctg gctgcacccg 2220
aatagatggg caggacattt cacaggtgac ccaggcctct ctccggtctc acattggagt 2280
tgtgccccaa gacactgtcc tctttaatga caccatcgcc gacaatatcc gttacggccg 2340
tgtcacagct gggaatgatg aggtggaggc tgctgtctcag gctgcaggca tccatgatgc 2400
cattatggct ttccctgaag ggtacaggac acaggtgggc gagcggggac tgaagctgag 2460
cggcggggag aagcagcgcg tcgcattgc ccgcaccatc ctcaaggctc ctcaagctat 2520
tctgctggat gaggcaacgt cagcgttgga tacatctaat gagagggcca tccaggcttc 2580
tctggccaaa gctctgtcca accgcaccac catcgtagtg gcacacaggc tctcaactgt 2640
gggtcaatgc gaccagatcc tgcgtcatca ggatggctgc atcgtggaga ggggacgaca 2700
cgagggtctg ttgtcccgag gtgggggtgt tgctgacatg tggcagctgc agcagggaca 2760
ggaagaaacc tctgaagaca ctaagcctca gccatggaa cggtgacaaa agtttggcca 2820
cttccctctc aaagactaac ccagaaggga ataagatgtg tctcctttcc ctggcttatt 2880
tcactctggt ctctgggggt ggtgtagct aaagggaact tccgaaaaaa 2940
catcttttgg ggaataaaaa atgtggactg tgaaaaaaa aaaaaaaa aaa 2993

```

<210> 188

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005689

<400> 188

ggaaaggagc ctttccgaaa aacatctttt ggggaaataa aaatgtggac tgtgaaaaaa 60

<210> 189

<211> 1830

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005749

<400> 189

```

ggggagtgtg aacctaatTT tgtggcgtag cagctatgca gcttgaaaTC caagtagcac 60
taaaatTTtat tatttctgat ttgtacaata agcttcccgag gagcgtgtc aacatttttg 120
gtgaagaact tgaagaactt cttaagaaga aatatgaagg gcactggtat cctgaaaagc 180
catacaaagg atcgggggtt agatgtatac acatagggga gaaagtggac ccagtgtatt 240
aacaagatcat caaagagagt ggtttggaca ttgatgtat tcgtggcaat ctgccacagg 300
atctttagtgt ttggatgcac ccaatttgagg ttcttaacca aatttgttga aagggaccag 360
tgaaggtgct ttacgtggat gataataatg aaaaatgagt tgaagtgttat aaggagatca 420

```

```

aaaaacagctt taaccagag gccagggtt ttatgccc atgtgaccca gctcatcag 480
tgtccagctc tccatcgct cctttgggt actctgctgc tgaagccct acctcatgc 540
ccggctccac tcagccttta acctttacca ctggcacttt tgetgccc aagttcggt 600
ctacccaaa gaagcaatagt ggcgtagca acaagggtgc acgtactct cccatcaacc 660
tcggcttgaa tgtgaatgac ctcttgaagc agaaagccat ctcttccca atgcactctc 720
tgtatgggct tggcttgggt agccagcagc acagcagcag cagcggcagc 780
gccaccggcc accaccacca ccacagcagc aacaacagca gaaaacctct gctctttctc 840
ctaagcccaa ggaatttatt ttctctaata tgcagggtca aggtagtatg accaatggaa 900
tgttccagag tcagagcccc cttaacctca gtcctctcca gtacagtaat gcttttgtatg 960
tgtttgcagc ctatggaggc ctcaatgaga agtcttttgt agatggctgt aattttatgt 1020
taaataacat gcagtattct aaccagcaat tccagcctgt tatggctaac taaaaaaaag 1080
aaaatgtatc gtacaagtta aaatgcacgg gcccaagggt gatttttttt tccacctctc 1140
tgagaatttt tttttttaag ctatagataa ggatacattc aagcttgggt aaaaaataa 1200
taataaaaaa tgcattcttt ttcatttgcc aaccaagcac aaagtattt tatactgact 1260
gtatatttta aagtatctc tcagatatgg cctcttacg tatttaagat atagcaagga 1320
catggctgat ttttttttat aaaaattggc actaataagt gggtttatgt gtcctttcta 1380
attgtataat ttaatttagt acaagtttg taaaataatc gaggatata atatatgtt 1440
ctcacgacat ggtattgcat ttatatcttt taactacagt gatctgtgac agcagcagct 1500
tcagtgtgta ttttttttac tgaaattgta aaataccat cttaaaagca ctaactattc 1560
aaaaattgtg gtacaggata ttctttagt ggtggaatta aaatgtacga atactgtctt 1620
ttcaaaaaaa atgtattttc tgttaaaagt ttaaaagatt ttgctatata ttatgaaaga 1680
aaaatgtaat cgtaaatatt aattttgtac ctatattgtg caatacttga aaaaacggt 1740
ataaagatat ttgagtcag tgtctacat gtttaagggt actgaaatag tttatattaa 1800
gtttgtatta aaattcttta aaataaaaa 1830

```

<210> 190

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005749

<400> 190

aaacctctgc tctttctcct aatgccaaagg aatttttttt tcctaataatg caggggtcaag 60

<210> 191

<211> 1534

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005804

<400> 191

```

ggaagcgcag caactcgtgt ctgagcgccc ggaggaaaac cgaagttgga agtgtctctt 60
agcagcgccc ggagaagaac ggggagccag catcatggca gaacaggatg tggaaaacga 120
tcttttggat tacgatgaag aggaagagcc ccaggctcct caagagagca caccagctcc 180
ccctaagaaa gacatcaagg gatcctacgt ttccatccac agctctggct tcgggacttt 240
tctgtcgaag ccggagctcc tgcgggccat cgtggactgt ggctttgagc atcctctcga 300
ggtcagcatg gagtgcattc ccaggcccat cctgggcatg gacgtcctgt gccaggccaa 360
gtccgggatg ggaagacag cgtctcttgt gctggccacc ctacagcaga ttgagctcgt 420
caacggacag gtgacggtcc tggctcatgt ccacagcagg gagctggcct tccagatcag 480
caaggatatg gagcctttt ccaagtacat gccacgcgtc aaggtgtctg tghttctcgg 540
tggctctctc atcaagaagg atgaagaagt gttgaagaag aactgtcccc atgtcgttgt 600
gggaaccccc ggcgcgcatcc tggcgctcgt gggcaatagg agcttcagcc taagaatgt 660
gaagcacttt gtgctggagc agtgtgacaa gatgctggag cagctgggga tgcgcgggga 720
tgtgcaggag atcttcgcc tcacaccaca cgagaagcag tgcattgatg tcagcgccac 780
cctgacaagc gacatccggc ctgtgtgcag gaagttcatg caggatccca tggaggtgtt 840
tgtggacgac gagaccaagc tcacgtgcea cggcctgcag cagtactacg tcaaaactcaa 900

```

```

agacagtgaag aagaaccgca agctctttga tctcttggat gtgctggagt ttaaccaggt 960
gataatcttc gtaacgtcaag tgcagcgctg catggccctg gccagctccc tegtggagca 1020
gaacttcccc gccatcgcca tccaccgggg catggcccgag gaggagcgcc tgtcacgcta 1080
tcagcagctc aaggatttcc agcggcggat cctgggtggcc accaatctgt ttggccgggg 1140
gatggacatc gagcgagtcac acatcgtctt taactacgac atgcctgagg actcgacac 1200
ctactcgacac cgggtggccc gggcggtcg ctttggcacc aaaggcctag ccatcacttt 1260
tgtgtctgac gagaatgatg ccaaaatcct caatgacgtc caggaccggt ttgaagttaa 1320
tgtggcagaa cttccagagc aaatcgacat ctccacatac atcgagcaga ccgggttaacc 1380
accacgtgcc agcggcgccc accgggagcc gcccgcatgc agcttcacct cccctttcca 1440
ggcgccactg ttgagaagct agagattgta tgagaataaa cttgttatta tggaaaaaaa 1500
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1534

```

<210> 192

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005804

<400> 192

gttgagaagc tagagattgt atgagaataa agtgttatta tgaatgaag aagcctcacc 60

<210> 193

<211> 1416

<212> DNA

<213> Homo sapiens

<220>

<221> Modified_base

<222> 1 ... 1416

<223> n = a, c, g, or t

<300>

<308> NM_005945

<400> 193

```

aggaattccg gaattccgga attccgatgg atggaacaga aaataaatct aagtttgggtg 60
cgcaaccgat tctgggggtg tcccttgccg tctgcaaaagc tgggtgccgtt gagaagggggg 120
tcccctgtac cggccacatcg cgtacttggc tggcaacttc gaagctatcc tgccagctccc 180
ggcgttcaag tgtcatcatc aatggcggtt ctcatgtctg caacaagctg gccatgcaga 240
gtctgtccct ccagtcggtg cagcaaaactc agggaagcca tgcccgattg gaggcagaggt 300
ttaccacaac ctgaagaatg tcatcaagga gaaatatggg aaagatgcca ccaatgtggg 360
gatttgcgcg ggtttgtctc caacatcctg gagaataaag aaggcctgga gctgctgaag 420
actgctattg gaaagccctg cctacactgt aaaggtggtc atggcatgga cgtagcgccc 480
tcggagtctt tcaggttcagg gaactatgac ctggacttca agtctcccca tgaccgccag 540
aggtacatct cgcctgacca gctggctgac ctgtacaagt ccttcatcaa ggactacca 600
gtgggtctcta tcgaagatcc ctttgaccag gatgactggg gagcttcaga agttcacagc 660
cagtgacagga atccaaggtag tggggggatg actcacagtg accaaccocaa agaggatcgc 720
caaggcggtga acgagaagtc ctgcaactgc ctctgcctca aagtcaccaa gattggctcc 780
gtgacagagt cctttcagggc gtgcaagctg gcccaaggcca atggttgggg cgtcatgtgtg 840
tctcatcgct cggggggagac tgaagatacc ttcatcgctg acctggttgt gggcgctgtc 900
actggggcag atcaagactg gtgcgccctg ccgatcacgc gcttggccaa gtacaaccag 960
ctctccagaa ttgaagagga gctgggcagc aaggctaagt ttgccggcag gaacttcaga 1020
aacccttggc ccaagtaagc tgtgggcagg caagccttgc gtcaacctgt ggtctacagac 1080
ccctccctct gtgtcagctc aggcagctcg agggcccccga ccaacactgt cagggggtccc 1140
tgtagttaag cgcgccacgc cgtggagttc gtaccgcttc cttagaactc tacagaagcc 1200
aagctccctg gaagccctgt tggcagctct agcttctgag ttgtgtaatt ggcccaagtc 1260
attgtttttc tgccttactc ttcccacaa gtcttagagt catgtgagcc tngtgtcatc 1320
tcgggggtgg ccaacaggcta gatcccggt ggttttggc tcaaaataaa aagcctcagt 1380
gaccatgtaa aaaaaaaaaa gaattccgga attccg 1416

```

<210> 194
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005945

<400> 194
 ttgtgtaatt ggcccaagtc attgtttttc tcgccttact ttccaccaag tgtctagagt 60

<210> 195
 <211> 961
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006014

<400> 195
 ggcgaccacg gtgtcttcaa aagccccgtc agggttggct tcctggggcc ggaccgactg 60
 tgggtcagtt tgcaccacgc ctctggaatc gagttacgcg cgaaaggcca gattttctgg 120
 aggaaaacgc agcctctcaa ccgctgacgc ggtctcagaa ggcccccgcc agggccgctt 180
 gcggggaact gaccacgcgc cagtcaggct ctccagggac ctgcgcaggc gcgtgtgggc 240
 ggagtcgtgc gcagggggag gggcttcggg aaggagccac agagagggag gggcgtagga 300
 cctgcgcttc ggggttgag tcggagcgc gcggcgccgc tcatgcggga cgcggatgca 360
 gacgcaggcg gaggcgctga cggcgggat ggccgggggt gccacagctg ccgcgggggc 420
 gtggacacag ccgcagctcc ggcgggtgga gctccccag cgacagcgcc aggtccgggc 480
 agagacgcgc cgtctgcggc cagggggtca cgaatgcggc cgcacatatt caccctcagc 540
 gtgcctttcc cgacccccct ggaggcgaaa atcgcccatg ggtccctggc accagatgcc 600
 gagccccacc aaagggtggt tgggaaggat ctacacagtga gtggcaggat cctggtgtgc 660
 cgctggaag ctgaagactg tcgcctgctc cgaatttccg tcatcaactt tcttgaccag 720
 ctttccctgg tggtcgggac catgcagcgc ttggggccc cagt ttcccg ctaagcctgg 780
 cctgggcaaa ttgagcgagg tcccactttg cgtctccttg taggcagtgc gtccatcctt 840
 ccctagggca ggaattccca cagttgtctac ttctctggga gggcctcatg ttttatctgg 900
 ttcttaaatg tttgttacta cagaaaataa aactgcgcta ctaaaaaaa aaaaaaaaaa 960
 a 961

<210> 196
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006014

<400> 196
 ggccctcatgt tttatctggt tcttaaatgt ttgttactac agaaaaataa actgaggtat 60

<210> 197
 <211> 1648
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006086

<400> 197
 atgccccgaga tcgtgcacat ccaggccgcg cagtgcggca accagatcgg ggccaagttc 60
 tgggaagtca tcagtgatga gcatggcctc gacccccagc gcaactacgt gggcgactcg 120

```

gacttgacgc tggagcggat cagcgtctac tacaacgagg cctctttctc caagtacgtg 180
cctcgagcca ttctgggtga cctggaaccc ggaaccatgg acagtggtcc ctcaggggcc 240
tttggacatc ttctcagggc tgacaatttc atcttttggt agagtggggc cggaacaac 300
tgggccaaag gtccatcac ggagggggcg gagctgggtg attcggtctt ggatgtggtg 360
gggaagaggt gtgaaactg cgactgcctg cagggtcttc agtgaccca ctgcgtgggg 420
ggggggagcg cctccggcat gggcacgttg ctcatcagca aggtgcgtga ggagtatccc 480
gaccgcatac tgaacacctt cagcgtcgtg cctcaccca agtggtcaga cagcgtgggt 540
gaacctatac acgcacagct gtccatccac cagctgggtg aaaacacgga tgaacctac 600
tgcatcgaca acgagggcgt ctacgacatc tgcttccgca cctcaagct ggccacgccc 660
acctacgggg acctcaacca cctggtatcg gccaccatga gcggagtac cactccttg 720
cgcttccggg gccagctcaa cgctgacctg cgcaagctgg ccgtcaacat ggtgcccttc 780
ccgcgcgtgc actttctcat gcccggtctc gcccccctca ccaggcgggg cagccagcag 840
taccggggcc tgaacctgcc cgagctcacc cagcagatgt tcgatgccaa gaacatgatg 900
ggcgccgtgc acccgcgcca cggcccgctac ctgacgggtg ccaccgtgtt ccggggcgcg 960
atgtccatga aggaggtgga cgagcagatg ctggccatcc agagcaagaa cagcagctac 1020
ttcgtggagt ggatcccca caactgtgaag gtggccgtgt gtgacatccc gcccccgcg 1080
ctcaagatgt cctccacctt catcggggaa agcacggcca tccaggagct gtccaagctg 1140
atctccgagc agttcacggc catgttccgg cgcaaggcct tctctgacct gtacacgggg 1200
gagggcatgg acgagatgga gttcaccgag gccagagaca acatgaacga cctggtgtcc 1260
gagtcaccgc agtaccagga cgccacggcc gaggaagagg cgcagatgta cggaagcagc 1320
gaggaagggt cggaggccca gggccccaag tgaactgct cgacagctga gtgagagcta 1380
ggtggcgccc gggggcgaag ccagcagtg cttaaaccctc ggagccatct tgctgcgcag 1440
acctgccttt ccccatcgcc ctagggtctc ctgtgcgcct tcttcagta ttatggcct 1500
cgctctcccc caactaggcc acgtgtgagc tgctcctgtc tctgtcttat tgcagctcca 1560
ggctgaagct tttaagggtt ttgtttttac tggttgtgtg ttatattttc ggggatactt 1620
aataaactta ttgtgtcag ataccctt 1648

```

<210> 198

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006086

<400> 198

```

tttttactcg tttgtgttta tatttttcgg gatacttaat aaatctattg ctgtcagata 60

```

<210> 199

<211> 3074

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006096

<400> 199

```

aacaacacct gcctggctcc cagctggtgc tgaagctcgt cagttcacca tccgccctcg 60
gcttcgcggg ggctgtgggc cgcagcctcc ggcacgctcc ttctctttct ccttcggtt 120
agccagagtg cagcagggac atgtctcggg agatgcagga tgtagacctc gctgagtgta 180
agcctttggt ggagaaaggg gagaccatca ccgctcctc gcaagagt tt gatgtccagg 240
agcaggacat cgagacttta catggctctg ttacgctcac gctgtgtggg acctccaagg 300
gaaacccggc ttctcatctc acctaccatg acatcgccat gaaccacaaa cctgtctaca 360
accctctctt caactacgag gacatgcagg agatcaccca gcactttgcc gtctgcgcag 420
tggagccccc tggccagcag gacggcgca cctccttccc cgacgggtac atgtacctcc 480
ccatggatca gctggctgaa atgcttctcg gagtcttcca acagtttggg ctgaaaagca 540
ttattggcat gggaacagga gcaggcgctc acatcctaac tcgatttgct ctatacaacc 600
ctgagatggt ggaagggcct tctctatata acgtgaacct ttgtcggaag gctggatgg 660
actggggcgc ctccaagatc tcaggatgga cccaagctct gccggacatg gtggtgtccc 720
acctttttgg gaaggagaa atgcagaatg acgtggaggt ggtccacacc taccgcagc 780
acatttgtaa tgacatgaac ccggcaaccc tgcacctgtt catcaatgcc tacaacagcc 840

```

```

ggcgcgacgt ggagattgag cgaccaatgc cgggaaccca cacagtcacc ctgcagtgcc 900
ctgctctggt ggtggttggg gacagctcgc ctgcagtgga tgccgtggtg gagtgcgaact 960
caaaattgga cccaacaaag accactctcc tcaagatggc ggactgtggc ggccctccgc 1020
agatctccca gccggccaag ctgcgtgagg ccttcaagta ctctgtgagc ggcatgggat 1080
acatgcctcc ggctagcatg accgcctcga tgcggtccgc cacagcctct gggtccagcg 1140
tcactttctt ggaaggcaacc cgcagccgct cccacaccaa cgagggcgcc gaagcccgct 1200
ccacaccaa cgagggcacc cgcagccgct cgcacaccaa cgagggggcc caacctggaca 1260
tcaccccaaa ctcggtgtgt gctgggaaca gcgccgggcc caagtccatg gaggtctcct 1320
ctcaggcgcc ctgcccaagt gccgccccc gactctgata tctgtatgtg cccctcctc 1380
ccggccctct ttctgccccc tgccctgccat actgcgcta actcggtatt aatccaaagc 1440
ttattttgta agagttagct ctggtggaga caaatgaggt ctattacgtg ggtgccctct 1500
ccaaaggcgg ggtggcggtg gaocaaagga aggaagcaag catctccga tcgatccctc 1560
ttcatttaac cagtggccgg ttgccactct cctcccctcc ctccagaca ccaaaactgcc 1620
aaaaacaaga cgcgtagcag cacacacttc acaagccaa gcctaggccg ccttgagcat 1680
ctggttcaaa acgggtgccc ggtcagaagg ccagccgcc acttccggtt tctcttttaa 1740
ctgaggagaa gctgatccag ttctccgaaa caaaactcct ttctcatatt gggagggggg 1800
taaatagtag actcaggcac ctcttttaaa caggcaaaac aggaaggggg aaaaggtggg 1860
attcatgtgc aggctagagg catttggaac acaaatcta cgtagttaac ttgaagaaac 1920
cgatttttaa agttggtgca tctagaaagc ttgtaatga agacttgatt 1980
ttctatgata cttctaattgt gcagcaaaag caggcgacaa aatctcctgg ctttacagac 2040
aaaaatattt cagcaaacgt tgggcatact ggtttttgaa gccttttagtt ctgtctctcg 2100
cctctccctcc acagcccaaa cctcccaccc ctgatatact agccagtgat tattcttgtt 2160
cagggagaag atcattttaga ttgtttttgc attccttaga atggaggcca acattccaca 2220
gctgcccctg ctgtgatgag tgtccttgca gggccggagg taggagcact ggggtggggg 2280
tggaattggg gttactcgat gtaagggtatt cctgttgtgt gtgttgagat caagtgcagt 2340
tgtgtattct gtgataccca gcttggttcc aggaattttg tgtgattggc ttaaatccag 2400
ttttcaatct tcgacagctg ggcctggaacg tgaactcagt agctgaaact gctcgaccgc 2460
gtcacgttct tggatcctca gaactcttgg ctctgttcgg ggtgggggtg ggaactcag 2520
tggggagcgg tggctgagaa aatgtaagga tcttggaaata catattccat gggactttcc 2580
ttccctctcc ttgcttccct ttctcgtctc cctcaacttt cgcgaaatgg gacgcacca 2640
ctgacgtttc tggggcgcca gtgcggtctc caggttctct tactactgcc ttgtactttt 2700
cattttggct caccgtggat ttctcatag gaagtttggc cagagtgaat tgaatattgt 2760
aagtacgcca ctgggacccc aggtattctg ggaccccgca gttggaggga ggaagtatgc 2820
cagccttcca ggtggcgtga gaggcaatga ctcgttacct gccgccatc accttggagg 2880
ccttccctgg ccttgagttag aaaagtcggg gatcggggca agagagcgct agtacggatg 2940
ggaaactatt gtgcacaaag cttccagag gagtttctta atgatatatt tgtatttatt 3000
tccagaccaa taattttgta actttgcagc ggaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3060
aaaaaiaaaa aaaa 3074

```

<210> 200
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006096

<400> 200
gagtagcggat gggaaactat tgtgcacaag tctttccaga ggagtttctt aatgagatat 60

<210> 201
<211> 2148
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006115

<400> 201
gcttcagggt acagctcccc cgcagccaga agccgggccc gcagcgcttc agcacccgtc 60
cgggacaccc caccgcgttc ccaggcgtga cctgtcaaca gcaacttcgc ggtgtggtga 120

```

actctctgag gaaaaaccat tttgattatt actctcagac gtgcgtggca acaagtgact 180
gagacctaga aatccaagcg ttggagggtcc tgaggccaag ctaagtgcgt tcaaaatgga 240
acgaagcgct ttgtgggggtt ccattcagag ccgatacatc agcatgagtg tgtggacaag 300
ccacgggaga cttgtggagc tggcagggca gagcctgctg aaggatgagg cctggccat 360
tgccgcctgt gagttgtctgc ccagggaagct ctctccgcga ctcttcattg cagcctttga 420
cgggagacac agccaagacc tgaaggcaat ggtgcaggcc tggcccttca cctgctctcc 480
tctgggagtg ctgatgaagg gacaacatct tcacctggag accttcaaa cttgtcttga 540
tggacttgat gtgctccttg cccaggaggt tcgcccagg aggtggaaac ttcaagtgt 600
ggatttaccg aagaactctc atcaggactt ctggactgta tggctctgaa acagggccag 660
tctgtactca ttctcagagc cagaagcagc tcagcccatg acaaagaagc gaaaagtga 720
tggtttgagc acagaggcag agcagccctt cattccagta gaggtgctcg tagacctgt 780
cctcaaggaa ggtgcctgtg atgaattggt ctctacctc attgagaag tgaagcgaaa 840
gaaaaatgta ctacgcctgt gctgtaagaa gctgaagatt tttgcaatgc ccatgcagga 900
tatcaagatg atctgaaaa tgggtgcagct ggactctatt gaagattttg aagtgaactg 960
taacctggaag ctacccacct tggcgaaatt ttctccttac ctgggcccga tgattaatct 1020
cgctgagact cctcctcccc acatccatgc atctctctac atttcccgg agaagggaag 1080
gcagtatatc gccacgttca cctctcagtt cctcagttcty cagtgcctgc aggtctctta 1140
ttgggacctt ttatttttcc tttagggccg cctggatcag ttgctcaggg acgtgatgaa 1200
cccttggaa accctctcaa taactaactg ccggttttcg gaaaggggat tgatgatct 1260
gtccagagtg cccagcgtca gtcagctaag gtgctgagtg ctaagtgggg tcatgctgac 1320
cgatgaagt cccgagcccc tccaagctct gctggagaga gcctctgcca cctccagga 1380
cctgggtctt gatgagtggt ggtatccgga tgatcagctc cttgccctcc tgccctccct 1440
gagccactgc tcccagctta caaccttaag cttctacggg aattccatct ccatatctgc 1500
cttgacagtg ctctcgcagc acctcatcgg gctgagcaat ctgaccacg tgctgtatcc 1560
tgtccccctg gagagttag aggacatcca tggtaacctc caoctggaga ggcttgactc 1620
cttgcagtcg agcgtcaggg agttgctgtg tgaattgggg cggcccagca tggctggctg 1680
tagtgccaac cctgtctctc actgtgggga cagaaccttc tatgaccgg cagccatctc 1740
gtgcccctgt ttcattgccta actagctggg tgcacatate aaatgtctta ttctgcatac 1800
ttggaaacta aagccaagat gtgcatgcat cttgaagcaa caaagcagcc acagtttccag 1860
acaaatgttc agtgtgagtg aggaatacat gtaagtgag gaaaaaacat tcagacaatat 1920
gttcagtgag gaaaaaaagg ggaagttggg gataggcaga tggttgactg aggaagtaat 1980
gtgatctttg gggagataca tcttatagat ttgaaaaatg aatctgaatt tctaaagggg 2040
gattctggct tgggaagtac atgtaggagt taatccctgt gtgactgttt gtaagaagaa 2100
tgttgaaaat aagagaagc aatgtgaagc aaaaaaaa 2148

```

<210> 202
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006115

<400> 202
tgaggagata catcttatag agttagaat agaactctgaa ttctctaaag gagattctgg 60

<210> 203
<211> 1051
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006332

<400> 203
ggaccgcgcg cgtgttaaaag ggcgttattt cccaggcagc cgtgcagtc gccacacctt 60
tgccctgtct gcgatgaccc tgtcgccact tctgctgttc ctgccacgc tgctgctgct 120
gtgcagcgtc cccaagcgcg oggtgaagc cctccctctg caagcgttag acttctttgg 180
gaatgggcca ccagtttaact acaagacagg caatctatac ctgcggggcg cctgaagaa 240
gtccaatgca ccgctgttca atgtgacctt ctactatgaa gcactgtcgc gtgctgcgcg 300


```

agccttctgt atccggggagc tcttcccaac atggctgttg gtcattggaga tctcaatgt 360
cacgtctgtg cctacaggaa acgcacagga acaaaatgtc agtggccagt gggagttcaa 420
gtgcacagcat ggagaagagg agtgcaaat tcttgacatg tgtctgcatg gaggcctagc 480
acttgacatg gaggtagcct tcttgacatg tgtctgcatg gaagagtttg aggcattgga 540
gaagaagtctg ccaatgatgcc tgcagctcta cgcgccaggg ctgtcgccag acactatcat 600
ggagtgctga atgggggacc gcggcatgca gctcatgcac gccaacgccg agcgggacaga 660
tgctctccag ccaccacacg agtatgtgcc ctgggtcacc gtcaatggga aaccttggga 720
agatcagacc cagctcctta cctctgtctg ccagttgtac caggggcaaga agccggatgt 780
ctgcctcttc tcaaccagct cctcaggag tgtttgcttc aagtgtatggc cgggtgagctg 840
cggagagctc atgggaagcg agtgggaaacc cggctgcctg cctttttttc tgatccagac 900
cctcgccacc tgctacttac caactggaaa attttatgca tcccatgaag ccagatatac 960
caaaattcca ccccatgac aagaatcctg ctccactaag aatggtgcta aagtaaaact 1020
agtttaataa gcaaaaaaaa aaaaaaaaaa a 1051

```

<210> 204
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006332

<400> 204
aaattccacc cctgatgtaa gaatcctgct ccactaagaa tgggtgctaa gtaaaactag 60

<210> 205
<211> 1714
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006417

```

<400> 205
ggggcatttt gtgcctgcct agctatccag acagagcagc tacctcagc tctagctgat 60
actacagaca gtacaacaga tcaagaagta tggcagtgac aactcgtttg acacggttgc 120
acgaaaagat cctgcataat cattttggag ggaagcggct tagccttctc tataagggtta 180
gtgtccatgg attccgtaat ggagttttgc ttgacagatg ttgtaataca gggcctactc 240
taacagtgat ttatagttaa gatcatatta ttggagcata tgcgggaagag agttaaccagg 300
aaggaagata tgcctccatc atcctttttg cacttcaaga tactaaaatt tcgaattgga 360
aactaggact atgtacacca gaaacactgt tttgttggta tgttacaata tataactccc 420
caactaattt ccagatgatg ggaagaataa gaaaagtgtat tatggactta aagaccaatg 480
aaaatcttgg acttgctcaa aattgtacta tctctattca ggattatgaa gtttttcgat 540
gcgaagattc actggtatgaa agaagaataa aagggtgcct tgagctcagg aagagcttac 600
tgtctgcctt gagaaactat gaaccataat gatccctggc tcaacaaata cgaattctgc 660
tgtctggctt aattggagct ggggaagcca gctttttcaa ctcatgtagg tctgttttcc 720
aaggcattgt aacgcattcag gctttgggtg gcactaatat aactgggata tctgaaagt 780
ataggacata ctctattaga gaagggaag atggcaataa cctgccgttt attctgtgtg 840
actcactggg gctgagtgag aaagaaggcg gcctgtgcag ggaatgaata ttctatatct 900
tgaacggtaa cattcgtgat agataccagt ttaatcccat ggaatcaatc aaattaaatc 960
atcatgacta ccttgattcc ccactcgtga aggcacagat tcatgtgtgt gcaatttgtat 1020
ttgatgccag ctctattcaa tacttctctc ctcatgatgt agtaagatc aaaagaattc 1080
gaaggagatt ggtaaacgct ggtgtggtac atgtggcttt gctcactcat gtgtagaga 1140
tggtattgat tacaaaaggt gaccttatag aaatagagag atgtgagcct gtgaggtcca 1200
agctagagga agtccaaaga aaacttggtat ttgctcttct tgacatctcg gtgttagaga 1260
attattccctc tgagtgggag ctggaccctg taaaggatgt tctaattctt tctgctctga 1320
gaacaaatgt atgggctgca gatgaactct tagaggatgt gccttttgag caaataggga 1380
atctaaggga ggaattatc aactgtgcac aaggaaaaaa atagatatgt gaagggttca 1440
cgtaaaattc ctcatcatc acgaagataa aatcagaaa ggagaaaaac cagacccaag 1500
agaagtatct aagaccaaag ggaatgtgtt tattaatgtc taggatgaag aagtcgatag 1560
aacatttgat tacttgtaaa taactagaaa taacatgatt tagtcaatat tgtgaaaaat 1620

```

agtaataatt ttctctggat ttatgttctg tatctgtgaa aaaaataaatt tcttataaaa 1680
ctcggaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1714

<210> 206
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006417

<400> 206
atgacatatt ctatatcttg aacggttaaca ttctgtgatag ataccagttt aatcccatgg 60

<210> 207
<211> 3791
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006461

<400> 207
acagacggcg ggtgaacatg gcgtcctcga cttgggtctga gacgtgatag gccctgccttc 60
tggttgaaga tgtgcgaggt gaaaaaaact agcctcagcc tgtcgccttc gccccagacg 120
ggaaaaacat ctatgagaac tcctctccgt gaacttaccc tgcagcccggt tgcctccacc 180
acctctggaa aaagatcccc cgcttgctcc tcgctgaccc catcaactgtg caagctgggg 240
ctgcaggaag gcagcaacaa ctctgtctcca ttggattttg taaataacaa gaggacagac 300
ttatcttccg aacatttccg tcattctctca aagtggctag aaactgttca gcatgaatca 360
gatgagcagc ctctagatcc aatcccccaa attagcteta ctcctaaaac gtctgaggaa 420
gcagtagacc cactggggcaa ttatatgggt aaaaaccatcg tccctgtacc atctccactc 480
gggcagcaac aagacatgat atttgaggcc cgcttagata ccatggcaga gacaaaacagc 540
atatctttaa atggaccttt gagaacagac gatctggtga gagaggaggt ggcacccctgc 600
atggggagaca ggttttcaga agttgctgct gtatctgaga aacctatctt tcaggaaatct 660
ccgtcccatc tcttagagga gtctccacca aatccctggt ctgaacaact acattgtccc 720
aaggaaaacc tgagcagtag aactgaggct gtggtgtagg acttagtacc ttctgaaagt 780
aacgcctctt tgcctctctc tgttctctcg ctttccctct caactgcctt ggcagcagat 840
ttcgtgtcga atcatgttga ccagaggag gaaattgtag agcatggagc tatggaggaa 900
agagaaatga ggtttccacc acatctcaag gactctgaaa cagaagatca agcacttgtc 960
tcaagttgtg aagatatctt gtccacatgc ctgacaccaa atctagtga attggaatcc 1020
cagaagaatc cagcccccagc agtgaagaat gttggttaga ttcttgcttc tgatacagag 1080
tcttgatagt cccactggcc ctggctggaa aaaggtgtaa atacctccgt catgtctggaa 1140
aatctccgcc aaagcttacc ccttccctcg atgcttcggg atgtctcaat tggcactacc 1200
ccttctctca cttgtctggt ggggacttgg ttactcctt cagcaccaca ggaagaagat 1260
acaaacacat ccacagacag cctggttggc accaagcaca gacttcttga gacagacag 1320
ctcctgtgtg gccggccctcc agatctgact gccctgtctc gcatgactt ggaagataac 1380
ctgctgagct cctctgtcat tgtggagttt gctctcccgcc agctctcgga ctggaagagc 1440
cagctggctg tccttcaccc agaaaaccag gacagttaga cacagactga cacactcac 1500
agtgggataa ctaataaact tcagcatctt aaggagagcc atgagatgg acagccctca 1560
cagcagacca gaaatgtcat gcaactatcg gtgcttatct ctaaaagact gatattcctg 1620
cttcaactat cctctgttga tttagaagaa gataagacta ctgtgaatca ggaagtctcg 1680
cgctgcagaa cattgtctg ttgctgtttt gatttgcgtga agaaattgag ggcaagctc 1740
cagagctcca aagcagaaaag ggagaggcca aggcacagag aggaatggc tctcagagcc 1800
aaggatcgcg cagagatagt gttggaggct ttctgtgcac acgccagcca gcgcactcagc 1860
cagctggaaac agagcctagc atccatgcgg gaattcagag gcccttctgaa ggaatgccag 1920
acccaactcg tagggcttca tgccaagcaa gaagagctgg ttcagcagac agtgagtctt 1980
acttctacct tgaacaacga ctggaggctc attatacaac attgacagct ttggacagct 2040
tctgtgagtc ggtcccgaca actcacagag aaactcacag tcaagagcca gcaagccctg 2100
caggaaactg atgtggcaat tgagggaaaag caggaggttt caaggagctt ggaacaagtc 2160
tctgccagtc tagaggagtg caaaggccaa acagaaacac tggagttgga aaactctgt 2220
ctagcaacag atctccgggc tcaagttcag attctggcca acatggagac ccagctaaaa 2280

gagctacaga	gtcagcatca	ccattgtgcc	caggacctgg	ctatgaagga	tgagtacttc	2340
tgccagctta	cccagagcaa	tgaggagcag	gctgctcaat	gcgtaaagga	agagatggca	2400
ctaaaaacaca	tcgaggcgaga	actgcagcag	caacaagctg	tcctgggccaa	agaggtgcgg	2460
gacctgaag	agaccttgga	gtttgcagac	caggagaate	aggttgctca	cctggagctg	2520
ggtcaggttg	agtgtcaatt	gaaaaccaca	ctggaagtgc	tcggggagcg	cagcttgacg	2580
tgtgagaacc	tcaaggacac	tgtagagaac	ctaacggcta	aactggccag	caccaataga	2640
gataaccagg	agcaagatct	ggagaaaaca	cggcagtaact	ctcaaaagct	aggctcgtc	2700
actgagacaac	tacagagcct	gactctcttt	ctacagacaa	aactaaagga	gaagactgaa	2760
caagagaccc	tctctgtgag	tacagcctgt	cctccccccc	aggaaacccc	tctgcctaact	2820
gacaggacct	tctctgggaag	catcttgaca	gcagtggcag	atgaagagcc	agaactcaact	2880
cctgtgccct	tgcttggaag	tgacaagagt	gctttcaccc	gagtagcatc	aatggtttcc	2940
cttcagcccg	cagagaccoc	aggcatggag	gagagcctgg	cagaaatgag	tattatgact	3000
actgagcttc	agagtctttg	ttccctgcta	caagagtcta	aagaagaagc	catcaggact	3060
ctgcagcgaa	aaattttgtga	gctgcgaagct	aggctgcagg	cccagggaaga	acagcatcac	3120
gaagtccaga	aggcaaaaga	agcagacata	gagaagctga	accaggcctt	gtgcttgccg	3180
tacaagaatg	aaaaggagct	ccagggaagt	atacagcaga	atgagaagat	cctagaacag	3240
atagacaaga	gtggcgagct	cataagcctt	agagaggagg	tgacccacct	taccgcgtca	3300
cttcggcgct	cggagacaga	gaccaaaagt	ctccaggagg	cctggcaggg	cagctggact	3360
ccaaactgcc	gctattggcc	accaattgga	tcctaggaga	agtgtggctc	tctcaggagg	3420
tggaacaaact	gagagtgatg	ttcctggaga	tgaaaaaatga	gaaggaaaaac	tctctgatcaa	3480
gttcacagac	ccatagaagt	atcctagagg	agaaaccttc	gcgctctgac	aaggagttag	3540
aaaaactaga	tgacattggt	cagcatattt	ataagaccct	gctctctatt	ccagaggtgg	3600
tggaaggatg	caagaacta	cagggtatgc	tggaatttct	gagctaagaa	actgaagacc	3660
agaatttgtt	tcacctcttt	ttacctgcaa	taacccctta	ccccaatacc	aagaccaact	3720
ggcatagagc	caactgagat	aaatgctatt	taaataaagt	gtattttaatg	aaaaaaaaa	3780
aaaaaaaaa	a	3791				

<210> 208

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006461

<400> 208

ctgacaagga gttagaaaaa ctatgatgaca ttgttcacga tatttataag accctgctct 60

<210> 209

<211> 2856

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006516

<400> 209

tagtcggcgg	tcctccgagt	agcacgccag	ggagcaggag	accaaaacgac	gggggtcggg	60
gtcagagtcg	cagtgggagt	cccgggaccg	gagcacgagc	ctgagcggga	gagcgcgcgt	120
gcacgcggcg	tcgcaccocg	cgtaccocgg	gcagccagag	ccaccagcgc	agcgctgcc	180
tgagcccacg	cagcaaaag	ctgacgggtc	gcctcatgct	ggctgtggga	ggagcagtcg	240
ttggctccct	gcagtttgcc	tacaacactg	gagtcaccaa	tgccccccag	aaggtgatcg	300
aggagttcta	caaccagaca	tggttcaccc	gcatctctg	gagcatcctg	ccaccacgcg	360
tcaccacgct	ctggtccctc	tcagtggcca	tcctttctgt	tgggggcagc	attggctcct	420
tcctctgggg	ctctttcggt	aaccgctttg	gcggcgaggaa	tcacatgctg	atgatgaacc	480
tgctggcctt	cgtgtccgcc	gtgctcatag	gcttctcgaa	actgggcaag	tcctttgaga	540
tgctgatcct	gggcgcgctc	atcatcggtg	tgtaactcgg	cctgaccaca	ggcttcgctg	600
ccatgtatgt	gggtgaagtg	tcacccacag	cctttcgtgg	ggccctgggc	accctgcacc	660
agctgggcac	cgtcgtcggc	atcctcagcg	cccaggtgtt	cgccctggac	tcacatcag	720
gcacaagga	cctgtgcccc	ctgctgctga	gcacatcatt	catcccgccc	ctgctgcagt	780
gcacgtgct	gccctctctg	cccagagatc	cccgcttcc	gctcatcaac	cgcaacgagg	840

```

agaacccgggc caagagtgtg ctaaagaagc tgcgcgggac agctgacgtg acccatgacc 900
tgcaaggagat gaaggaagag agtcggcaga tgatcgggga gaagaaggtc accatcctgg 960
agctgtctccg ctcccccgcc tacccgcagc ccatcctcat cgctgtgggtg ctgcagctgt 1020
cccagacagct gtctggcatc aacgctgtct tctattactc cacgagcattc ttcgagaagg 1080
cgggggtgca gcagcctgtg tatgccacca ttggctccgg tatcgtcaac accgccttca 1140
ctgtcgtgttc gctgtttgtg gtggagcgag caggccggcg gacctgcac ctcataggcc 1200
tcgctggcatc ggcgggttgt gccatactca tgaccatcgc gctagcactg ctggagcagc 1260
tacccctggat gctctatctc agcatcgtgg ccatctttgg ctttgtggcc ttctttgaag 1320
tgggtctcgtg ccccatccca tggttcatcg tgggtgaact cttaagccag ggtccacgtc 1380
cagctgcctat tgcggttgca ggcttctcca actggacctc aaatttcatt ttctttgaag 1440
gcttcacgta ttgtggagcaa ctgtgtgggtc cctacgtctt catcatcttc actgtgctcc 1500
tggttctgtt ctccatcttc acctacttca aagttcctga gactaaaggc cggaccttcg 1560
atgagatcgc ttccggcttc cggcaggggg gagccagcca aagtgaataa acaccogagg 1620
agctgttcca tccctggggg gctgattccc aagtgtgagt cgcgccagat caccagcccg 1680
gcctgtctcc agcagcccta aggatctctc aggagcacag gcagctggat gagacttcca 1740
aacctgcagac atgtcagccg agccgggctc ggggctcctt tctccagcca gcaatgatgt 1800
ccagaagaat attcaggact taacggctcc aggtatttaa caaaagcaag actgtgtctc 1860
aaatctacat agacaagcaa caggttttat aaatttttta ttactgtatt tgttattttt 1920
atatcgacct gagtctcctg tgcccacatc ccaggcttca cctgaaatgg ttccatgcct 1980
gaggggtggag actaagccct gtcgagacac ttgcctctct caccagccta atctgtaggg 2040
ctggacctat gctctaagga cacactaatc gaactatgaa ctacaaagct tctatccag 2100
gaggtggctta tggccaccgg ttctgtctgg ctggatctcc ccactctagg ggtcaggctc 2160
cattaggatt tgccccttcc catctcttcc tacccaacca ctcaaattaa ctcttctctc 2220
cctgagacca gttgggagca ctggagtgca gggaggagag ggggaaggcc agtctgggct 2280
gccgggttctc agtctccttt gcactgaggg ccacactatt accatgagaa gagggctctg 2340
ggagcctctc aaactcaact ctaagaagaa catggagact cctgcccctg tgtgtataga 2400
tgcaaatatc ttatatatat caaatataa ttttgttgtt caatattaaa tacagacact aagtatatgt 2460
atatctggac aagccaactt gtaaaatcac cacctcactc tgaaatgctc ctgttacctc cctaaccaga 2520
tataaatggc tggtttttag aaacatgggt ttgaaatgctc tgtggattga gggtaggagg 2580
tttggatggg agtgagacag aagtgaagtg ggttgcaagg ctgtgcaagg cttagacctc 2640
gactcaggat ccagctccctt acacgtacct ctcatcagtg tctctctgtc caaaaattctg 2700
tttgatccct gtaccocaga gaatatatac attctttatc ttgacattca aggcatttct 2760
atcacatatt tgatagtttg tgttcaaaaa aacactagtt ttgtgccagg cgtgatgtct 2820
agcgttgaaa tcgcatttatt ttgaagtga agggaa 2856

```

<210> 210
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006516

<400> 210
 aaacagatat aaatggctgg tttttagaaa catggttttg aaatgcttgt ggattgaggg 60

<210> 211
 <211> 576
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006607

<400> 211
 atggctactc tgatctacgt tgaataaggaa attggagaac caggccaccg tgtggctgcc 60
 aaggatgtgc tgaagctgga gtctagacct tcaatcaaa caggatagg gatattccaa 120
 gttttaacac caogcttttg caaaacatac gatgtccatc cagccttacc taaagctacc 180
 agaaaggctt tgggcactgt caacagagct acagaaaagt cagttaagac caatggacc 240
 agaaaaaaca cagacacaa cttttctgac aaaaagatga ccgagaagac tgttaaaaa 300
 aaaaagtctg ttctctgcctc agatgacgcc tatccagaaa tagaaaaatt ctttcccttc 360

```

aatcttctag acttttagag ttttgacctg cctgaagagc gccagattgc acacctcccc 420
ttgagtgagg tgctctcat gatccttgat gaggaggagg agcttgaaaa gctgtttcag 480
ctgggccccct cttcacctgt gaaaatgcc tctccaccat gggaatgcaa tctgtttgca 540
gtctcttca agcattctgt cgaccctgga tgttga 576

```

```

<210> 212
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_006607

```

```

<400> 212
cgctatcca gaaatagaaa aattctttcc cttcaatctt ctgactttg agagttttga 60

```

```

<210> 213
<211> 2058
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_006820

```

```

<400> 213
gcacgaggaa gccacagatc tcttaagaac tttctgtctc caaacctggg ctgctcgata 60
aatcagacag aacagttaat cctcaattta agcctgatct aacccctaga aacagatata 120
gaacaatgga agtgacaaca agattgacat ggaatgatga aaatcatctg cgcaactgct 180
tggaatgttt tcttttagtc ttctctataa gtctagtgtt catggaggta gcattgga 240
tatgggtgaa agatgcagcc gtcagggtat tactataaca atggcttaca ttgattacaa 300
tatgattgta gcctttatgc ttggaataa tatataatta cgtgaaagtt ctacagagcc 360
aaatgattcc ctatggtttt cactcaaaa gaaaaatgac accactgaaa tagaaacttt 420
actctaaat acagacccaa aaattattga tgagcaactg gtgtgtcgtt tatcgaaac 480
ggatattttc attatatgtc gagataataa aatttatcta gataaaatga taacaagaaa 540
cttgaaacta aggttttatg gccaccgtca gtatttgaa tgtgaagttt ttcgagttga 600
aggaattaa gataacctag acgacataaa gaggataatt aaagccagag agcagagaaa 660
taggcctcta gcagacatca gagactatag gccctatgca gacttgggtt cagaaattcg 720
tattcttttg gtgggtccag ttgggtctgg aaagtccagt ttttcaatt cagtcaagtc 780
tatttttcat ggccatgtga ctggccaagc cgtagtgggg tctgatacca ccagcataac 840
cgagcgtgat aggatatatt ctgttaaaga tggaaaaaat ggaaatctc tgccatttat 900
gtgtgtgac actatggggc tagatggggc agaaggagca ggaactgtga tggatgacat 960
tcccacatc ttaaaagggt gtatgccaga cagatatcag ttaatttccc tgaacccaat 1020
tacacctgag cattctactt ttatcacctc tccatctctg aaggacagga ttcaactgtg 1080
ggcttatgtc ttagacatca actctattga caatctctac tctaaaatgt tggcaaaagt 1140
gaagcaaagt cacaaaagaa tattaaactg tggatatgca tatgtggcct tgcttactaa 1200
agtggtgatg tgcagtgagg ttcttcaaga caacttttta aacatgagta gatctatgac 1260
ttctcaaagc cgggtcatga atgtccataa aatgctaggg acttccattt ccaatatttt 1320
gatggttga aatatgcttt cagatttga actggacccc atgaaggata ttctcatctc 1380
ctgcactttg aggcagatgc tggggctctc agatgatttt tagaagatt ttgcctttga 1440
ggaaactggg gcaatttga gagggttaca gccctgcatt tgagataagt tgcccttatt 1500
ctgcactttg cgcocagctg tactggtgtg gccaatgag actcaatctc tatgacagc 1560
ctgtctcaga ttttgctttt gtctgttttg cctctctgct ttggaacagt catatctcaa 1620
gttcaaaagg caaacactga gaaagcgtgt gctcactctg gtccactctc aaacccccc 1680
tccattttc cgtaccattt acaattcagt ttctgtgaca tctttttaa cactggagg 1740
aaaaatgaga tattctctaa ttattctctc tataacactc tatatagagc tatgtgata 1800
ctaatacacat tgaataatag ttataaaatt attgtataga catctgtctc ttaaacagat 1860
tgtgtgtctt ttgagaaaca gcgtggattt tacttatctg tgatttcaga gagcttagca 1920
cagtgccctg taatagacaa gcatactgtc cattactttt ccttccact ctctccaa 1980
tcacattcac tttaaatttt tctgtatata gaaagggaaa ctgacctggg caacatgatg 2040
aaacccatc tccactgc 2058

```

<210> 214
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006820

<400> 214
tgagtctcttt gagaaacagc gtggatttta cttatctgtg tattcacaga gcttagcaca 60

<210> 215
<211> 2825
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006845

<400> 215
gcgaatttga ggtttcttgg tattgcgcgt ttctcttctt tgctgactct ccgaatggcc 60
atggactcgt cgcttcaggc cgcctctgtt cccggtctcg ctatcaagat ccaacgcagt 120
aatggttttaa ttacagtgcc caatgtaagg actgtgaact tggagaaatc ctgtgtttca 180
gtggaatggg cagaagaagg tgccacaaag ggcaaaagaga ttgattttga tgatgtggct 240
ccaataaaccc cagaactctt acagcttctt cccttacatc cgaaggacaa tctgcccttg 300
caggaaaaat taacaatcca gaacaaaaaa cggagatccg tcaactccaa aattctcgct 360
ccaagaagaaa gttctcgaag ccgctccact cgcctgtcca ctgtctcaga gcttcgcatt 420
acggctcagg aagaatgacat ggaggtggag ctgcctgcag ctgcaaaactc ccgcaagcag 480
ttttcagttc ctctcgcccc cactaggcct tccctgccct cagtggtgta aataccattg 540
aggatggtca gcgagagagat ggaagagcaa gtccattcca tccgtggcag ctctcttgca 600
aacctctgta actcagttcg gaggaatcca tgtcttgtga aggaagtgga aaaaatgaag 660
aacaagcgag aagaagaagaa ggcccagaac tctgaaatga gaatgaagag agctcaggag 720
tatgacagta gttttccaaa ctgggaattt gcccgaaatg ttaagaattt tcgggctact 780
ttggaatgtc atccacttac tatgactgat cctatcgaag agcacagaat atgtgtctgt 840
gttaggaaac gcccaactgaa taagcaagaa ttggccaaag aagaatttga tgtgatttcc 900
attctcagca agtgtctcct ctgtgtacat gaacccaagt tgaagtggga cttaacaaag 960
tatctggaga accaagcatt ctgctttgac ttgtcatttg atgaacagc tctgaattga 1020
gtgtctcaca ggttcacagc aaggccactg gtacagacaa tctttgaagg tggaaaagca 1080
acttgttttg catatggcca gacaggaagt ggcaagacac atactatgg cgagagacct 1140
ctctggaaag cccagaatgc atccaaaggg atctatgcca tggcctcccg ggaagctctc 1200
ctctgaaga atcaacctcg ctaccggaa ttgggcctgg aagtctatg gactcttctc 1260
gagattctaca atgggaagct gtttgacctg ctcaacaaga aggccaaact gcgcgtgctg 1320
gaggacggca agcaagcagg gcaagtgggt gggctgcagg agcatctggt taactctgct 1380
gatgatgtca tcaagatgct cgacatgggc agccctctga gaacctctgg ccagacattt 1440
gcccaactcca attcctcccg ctcccccagg tgcttccaaa ttattctctg agctaaaggg 1500
agaatgcgat gcaagtcttc ttgtgtagat ctggcaggga atgagcgagg ccgacacct 1560
tccagtctgt accgcgagac ccgcattggag ggcgcagaaa tcaacaagag tctcttagcc 1620
ctgaaggagt gcatcagggc cctgggacag aacaaggctc acaccccggt ccgtgagagc 1680
aagctgcacac aggtgctgag ggaactcttc attggggaga actctaggac ttgcatgatt 1740
gccacgatct caccaggcat aagctcctgt gaataactt taacaacct gagatatgca 1800
gacaggttca aggaagctgag ccccacagt gggccagtg gagagcatt gatccaaagt 1860
gaacacgaag agatggaagc ctgctctaac ggggcgctga ttcaggcgta ttatccaa 1920
gaagagagg aactgtcttc ccagatgtcc agctttaacg aagccatgac tcaatcagg 1980
gagctggagg aaagagctat ggaagagctc aaggagatca tacagcaagg accagactgg 2040
ctggaactct ctgatagatgac cgagcagcca gactatgacc tggagacctt tgtgaacaa 2100
cgggaatctc ctgtgcccga gcaagccaa ttttctcag cctcgagga tttctacaag 2160
gccttacgcct tggccatgca gctggaaag caggctagca gacaataaag cacgaagaaa 2220
cgccctcagg gacgaactgca aataaaaac ttgttgggtt gacacccag ctcttccctg 2280
ggcctcccca gagaactttg ggtacctggt gggctcagg agggctgag ctgggacagg 2340
ttctggtaaa tgccaagtat gggggcatct gggccaggg cagctgggga ggggctaga 2400

```

gtgacatggg acactccttt tctgttcctc agttgtcgcc ctcacgagag gaaggagctc 2460
ttagttaacc ttttgtgttg cctctctttc catcaagggg aatgtttctc gcatagagct 2520
ttctccgcag catcctgcct cgttggaact gctgctaatt gagagctccc tgggggtgtc 2580
ctggctctgg ggagagagac ggagccttta gtacagctat ctgctggctc taaacctttc 2640
acgcctttgg gccgagcact gaatgtcttg tactttaaaa aaatgtttct gagacctctt 2700
tctactttac tgtctcccta gagtccctaga ggaacctac tttatgtgtt 2760
tatacattgt atgtaacaat aaagagaaaa aataaaaaaa aaaaaaaaaa 2820
aaaaa 2825

```

```

<210> 216
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_006845

```

```

<400> 216
aaatgtttct gagacctctt tctactttac tgtctcccta gagtccctaga ggaacctac 60

```

```

<210> 217
<211> 823
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_007019

```

```

<400> 217
aaacgcgggc gggcgggccc gcagtcctgc agttgcagtc gtgttctcgg agttcctgtc 60
tctctgccaa ggcgcgcggg atggcttccc aaaaacgcga cccagcgcgc actagcgtcg 120
ccgcgcgcgg taaaggagct gagcogagcg gggcgcgccg ccggggctcg gtgggcacaaa 180
ggctacagca ggagctgatg accctcatga tgtctggcga taaagggatt tctgccttcc 240
ctgaatcaga caaccttttc aaatgggtag ggacctcca tggagcagct ggaacagtat 300
atgaagacct gaggtataag ctctcgctag agttcccagc tggctacctc tacaatgcgc 360
ccacagtgaa gttcctcagc cctcgctatc accccaacgt ggacacccag ggtaacatat 420
gcctggacat cctgaaggaa aagtgggtctg cctgtataga tgtcaggacc attctgctct 480
ccatccagag ccttctagga gaacccaaca ttgatagtc cttgaacaca catgctgccg 540
agctctggaa aaaccccaca gcttttaaga agtacctgca agaaacctac tcaagcagg 600
tcaccagcca ggagccctga cccaggctgc ccagcctgtc cttgtgtcgt ctttttaatt 660
tttctctaga tggctctgcc tttttgtgat ttctgtatag gaotctttat cttgagctgt 720
ggtatttttg ttttgttttt gtcttttaaa ttaagcctcg gttgagccct tgtatattaa 780
ataaatgcat ttttgccttt ttttagacaa aaaaaaaaaa aaa 823

```

```

<210> 218
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_007019

```

```

<400> 218
tggaataacc ccacagcttt taagaagtag ctgcaagaaa cctactcaaa gcaggtcacc 60

```

```

<210> 219
<211> 2831
<212> DNA
<213> Homo sapiens

```

```

<300>

```

<308> NM_007183

<400> 219

```

gaattccgga caggacgtga agatagttgg gtttgagggc ggccgccagg cccagggccc 60
gtggacctgc cgcacatgcag gacggttaact tctctgctgt ggccctgcag cctgaggccg 120
cgctgtgctc cctggcgctg cccctctgacc tgcagctgga ccgccggggc gccagggggc 180
cgagagccga gcggtctcgg gcagcccgcg tccaggagca ggtcccgccc cgcctcttgc 240
agctgggaca gcagcccgcg cacaaacgggg ccgctgagcc cgagcctgag cctacagactg 300
ccagagccag atccaggggg cagtaccaca cctcgaggc tggcttcagc tctcgctctc 360
agggcctcag tggggacaag acctcgggct tccggcccat cgccaaagcg cctacagactg 420
cagcctcctg gtctctccgc tccgcggtgg atctgagctg cagtccgagg ctgagttcag 480
cccacaatgg gggcagcgcc ttggggcccg ctgggtacgg ggtgcccag cccacccctc 540
ccatgcccac caggcccgctg tctctccatg agcgcggtgg ggttgggagc cgggcgcactg 600
atgacacact cctcctgcgc tcgctgcgcg tggggcccgg ggccctggac gaccgctaca 660
gcctgtgctg tgagcagctg gagcccgcgg ccacctccac ctacagggcc ttctcgtagc 720
agcgccaggc cagctccagc tccagccggg cagggggggt ggactggccc gaggccactg 780
aggtttcccc gaggccgacc atccgtgccc ctgcccgtgg gacccctgag cgattccaga 840
ctgagccagg gaggccgggg gttagggggg cagtgccggg ggccgtcctg gaggcagtg 900
ctcgagcgcc atctgtgcgc agcctcagcc tcagcctggc tgactcgggc cacctgcggg 960
acgtgcatgg gttcaacagg tacggtagcc accgaacctc gcagagactc agcagcggtt 1020
ttgtagacat tgacctgcgc tcagcagtca agtacctcat ggtctcagac cccaacctgc 1080
aggtgtcggg agccgacctac atccagcaca agtgctacag cgatgcagcc gccaaagaagc 1140
agggcccgag ccttcagacc gtgcttagcc ttgtgaagct cttcaaccac gccaaacctg 1200
aaagtgcagg ccatgccaca ggtgccatgc gcaacctcat ctacgacaac gctgacaaca 1260
agctgcggcc gtgtggaggag aacgggatct tcgagctgca cgggacactg cgggagcagg 1320
atgatgagct tcgcaaaaat gtccacagga tctctggaaa ccttctcatc agcgaccacc 1380
tgaaggaccc cctggccaga gacacgctgg agcagctcac ggacctgggt ttgagccccc 1440
tgtcgggggc tgggggtccc cccctcatcc agcagaagcg ctccgaggcg gagatctctc 1500
acaagcccac cggcttcttc aggaacctca gctcagctc tcaggccact cgccagaaga 1560
tgcggggagt ccaaggggctg gtggaagccc ttggtcaccct tatcaaccac gccctggagc 1620
cgggcaaatg cgaggaacaag agctgggaga acgcggtgtg cgtcctcggg aacctgtcct 1680
acggccctca cgcagagatg ccgcggtccg cgctgcagcg gctggagggt cgcggccgca 1740
gggacctggc gggggcgccg ccgggagagg tctgtgggctg cttcacgcgg cagagccggc 1800
ggctgcgcga gctgcccttc gcccgcgatg cgctcacctt cgcggaggtg tccaaggacc 1860
ccaagggcct caggtggctg tggagccccc agatcgtggg gctgtacaac cggctgctgc 1920
agcgtgcgga gctcaaccgg cacacgacgg aggcggccgc cggggcgctg cagaacatca 1980
cggcagcgga ccgcaggtgg gcgggggtgc tgagccgccc ggccctggag caggagcgtg 2040
ttctgaacct cctgctagac cgtgtcagga ccgcccagca ccaccagctg cctcactga 2100
ctggcctcat ccgaaaacctg tctcggaacg cttagaaca ggaagagctg tccacgaagg 2160
ttgtgagcca cctgatcgag aagctgccag gacagctggg tgagaagctg ccccccagcg 2220
aggtgctggc caacatcata gctgtgctca acaacctggt ggtggccagc cccatcgctg 2280
cccgagacct gctgtatttt gacggaactc gaaagctcat cttcatcaag aagaagcggg 2340
acagcccoga cagtgaaga gttcccccgg cagcatccag cctcctggcc aacctgtggc 2400
agtacaacaa gctccacgct gactttcggg cgaaggctca gacttctcgg 2460
gcccataggt gaagcctctc ggaggagaag gtgacgtggc ccagcgtcca agggacagac 2520
tcagctccag gctgcttggc agccagacct ggaggagaag gctaagtagc gaggggcccc 2580
tcgctggggc cctgtgtgac atctttgagg gtctctgggg accagagggg caggggtctt 2640
atagctgggg acttggtctc cgcaggcgag ggggtggggg agggctcaag gctgctctgg 2700
tgatgggggt ggtgacccag tcacattggc agaggtgggg gttggctcgt 2760
atcttgggat agccagacct ggggaataag atggccatga acagtcaaaa aaaaaaaa 2820
aaaaggaaat c 2831

```

<210> 220

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007183

<400> 220

ctggcagtat cttgggatag ccagcactgg gaataaagat ggccatgaac agtcacaaaa 60

<210> 221

<211> 2815

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007267

<400> 221

aggaagcgga ggaaggtgaa gtaggaccga attcctgtgc cgaagagggc tgcagtggga 60
 gagcaggatg ggggctccgg aggtggcgcc caggctctga gctaccctag gtctgcagac 120
 tagcgggcat tgccagaga catggcccag ccactggcct tcatcctoga tgtccctgag 180
 accccagggg accagggcca gggcccagc cctatgatg aaagcgaagt gcacgactcc 240
 ttccagcagc tcatccagga gcagagccag tgcaaggccc aggaggggct ggagctcgag 300
 cagagagcag ggaagtgac aggaagtagc cagcagacac tctggcgccc cgagggcacc 360
 cagagcacgg ccaactccg catcctggcc agcatgccca gccgcacat tggccgcagc 420
 cgaggtgccca tcatctccca gtactacaac cgacgggtgc agcttcggtg caggagcagc 480
 cgggccctgc tcgggaactt tgcctgcctc gctggcccca gcctccgctc gtacgacctg 540
 gagctggacc ccacggccct ggaggaggag gagaagcaga gcctcctggt gaaggagttc 600
 cagagcctgg cagtggcaca gcgggaccac atgcttcggg ggaatgccct aagcctggct 660
 gagaacacga gctcgcgaga gaagagcagg accccgaggg ggaagtggag ggcccaaggc 720
 ggcaagcggg gggctcgtct ctgctgtggc cggtccagat atgcctcgtc gctggccttg 780
 cacagcggcg gctcggcgct ctgacctccc ctgaggcccc tgatgcctgt gcgtcacgcc 840
 ctgaagcgca tcggggggcca gtctggctcc agcgtgctct cctacttctc ctttctcaag 900
 accctcgtgc ctttcaatgc cctcctgctg ctgctgctgg tggcctcat catgggcccct 960
 caggtgcgct tcccacccgc cctgccgggc cctgccccc gctgcacagg cctggagctc 1020
 ctcaacaggc cgggttgctt caccccacc gtcatgtact acggccacta cagttaacgc 1080
 acgctgaacc agccctgtgg cagccccctg gatggcagcc agtgacaccc cagggtggctt 1140
 ggctcgccct acaacatgac cctggcctac ctctccactg tgggctgtag cttctttatc 1200
 acctgataca cctcgtgtga cagcatggct cactctttcg gggagagcta ccgggtgggc 1260
 agcacctctg gcatccacgc catcacgcct tctgtcctct gggactacaa ggtgacgcag 1320
 aagcgggcct ccgcctccca gcaggacaat attgcaccc ggctgaagga gctcgtggcc 1380
 gagtggcagc tcgggcacag cccagggagc gtgtgcggga ggctgcgcca ggccgctgtg 1440
 ctgggggctg ttggtcgtct gtgtctgggg accgcgctgg gctgcgccgt ggccgtccac 1500
 gtcttctcgg agttcatgat ccagagtcca gaggctgctg gccaggaggg tgtgctgctg 1560
 gtctgcctcc tgggtggttg cctcctcaac ctggggggcc cctacctgtg cgtgtcctcg 1620
 gcgcctcgtg agccgcata ctccccgcta ctggagggtb acgtggccat ctcaggaaac 1680
 ctcatctcca actgggccat cctggggaca ctggtactac actggtcggg ccgagggtgt 1740
 ggcgtcctgc agggccagtg ctgggagagt ttttggggc aggaagctgt cgggttctcg 1800
 gtgtaggact tgcctcatat gtgtctggac acgctttttg ggaactggt ttggaagatt 1860
 attctccaga agaagctgaa gaggaggcgg aagccggagt ttgacattgc ccggaatgtc 1920
 ctggagctga tttatgggca gactctgacc ttgctggggg tgcctctctc gccctctctc 1980
 cccgcctgtg atgactcaaa gctgctgctc cgggcctcctg ctctctctatg tcaagaagac 2040
 gccactgccc agggcccgcg ccggccctgg ctggtctctat acatgagcac cgtctctctc 2100
 acgctgctct gcttcccgcg cctctgggg gcgcgtgtct tctcttgcta cgcgctgctg 2160
 caggtgaagc cctcgagcac ctgcggcccc ttccggaccc tggacacatc gtacgagccc 2220
 ggcaggggtg ggtgcgcga cctgaggggc gcaggcccca gggctcctcg ccgagggtgt 2280
 gtgcaccggt acctgaatga aacacacttc tttgtcttcc tgggtgcagc cctgctgctg 2340
 gccgtgatct acctcaacat ccaggtgggt cggggccagc gcaaggtcat cctgctgctc 2400
 aaggagcaga tcagcaatga ggttgaggac aaaaactctc taatcaacaa gcttcaactcc 2460
 attctacaga ggaaggagag gaggagagag agcagggttg ggaacaacga gaggctcgcc 2520
 gcacccctcg cctcgtccac agatgaacag gatgcctagg gggacggcga tggccctcac 2580
 gggcccgccc agcacctcga gaccacactg ttgctcccca gtgacctgc gctgacacca 2640
 ggcacaggaa gacagtcttg cctctcgaaa gccgcagctg cgctcaggtc ggagctgaaa 2700
 ggggtggtag atccgcttg ggcaccccca atgaaactgt cctcgtcctg gactctattt 2760
 attctgatta aagggttttt gcaaatggga aaaaaaaaaa aaaaaa 2815

<210> 222

<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_007267

<400> 222
ggtgaggaca aaatcttctt aatcaacaag cttcactcca tctacgagag gaaggagagg 60

<210> 223
<211> 1893
<212> DNA
<213> Homo sapiens

<300>
<308> NM_007274

<400> 223
atttaccggc ggcgggagag tgagggccca agtccgccct gctccgccac ttagccggcc 60
ccagacgctt ccctcggggc tgccaccggg tcgggcggcg ctgccggggc tagcggggcct 120
tcccgcaccc ggcgcggccc aaccgccacc gaacctctct gaagcggcgg ctgcctgggc 180
ccccacggcg ccagaatcgt acgcccgcgc gacgtctctg cagccttggc ggccctgggag 240
ggggggctcg ggggtggggc gggcgggggg cggggtcggc gcggggaggc cgcgttcgat 300
tcgccccggg cgcgcaggcc ccgcctcacc agccccatcg ctccacctct gccctcccc 360
tttatggcgc ggcccgggct cattcattcc gcgccgggccc tgccagacac ctgcgcctt 420
ctgcagccgc ccgcgcgcat cgcgcgcgca gcccccagca tgtcggggccc agacgtcgag 480
acgcgctcgc ccataccagat ctgcgggacg atgcggccag atgatgccaa cgtggccggc 540
aatgtccacg gggggacat cctgaagatg atcaggaggc caggcggcat catcagacc 600
cggcattgca acagccagaa cggggagcgc tgtgtggcgc ccctggctcg tgtcagcgc 660
accgacttcc tgttcccat gtgcacgggt gagggtggcg atgtcagcg gcagatcacc 720
tacacctcca agcactctgt ggaggtgcag gtcaacgtga tgtccgaaaa catctcaca 780
ggtgcacaaa agctgaccaa taaggccacc ctgtggtatg tgccccctgtc gctgaagaat 840
gtggacaagg tcctcggagt gcctcctgtt gtgtattccc ggccaggaca ggaggaggag 900
ggccggaagc ggtatgaagc ccagaagctg gacgcgatgg agaccaagtg gaggaacggg 960
gacatcgctc agccagtcct caaccacagc ccgaacactg tcagctacag ccagtcacg 1020
ttgatccacc tgggtgggccc ttcagactgc accctgcacg gctttgtgca cggaggtgtg 1080
accatgaagc tcattggatga ggtcgccggg atcgtggctg cagcgcactg caagaccaac 1140
atcgctcacg ctctcgttga cgccattaat ttcatgaca agatcagaaa aggcctgcgtc 1200
atcaccatct cgggacgcgt gaccttcacg agcaataagt ccattggagat cgaggtgttg 1260
gtggagccgc accctgttgc ggacagctct cagaagcgct accggggcgc cagtgccttc 1320
tcacacctacg tgtcgtctag ccaggaaggc aggtcgctgc ctgtgccccg gctgtgccc 1380
gaagccggag acgagaagaa gcgctttgag cgcgtcaaat ggocgtacct gcagatgaag 1440
gcgaagcgac agggccacgc ggagcctcag cctagactc cctcctcctg ccactgtgtc 1500
ctcagtagc cctggcaacg gcccagtgct coactcaact accgttcccc cccttggcca 1560
aaaaacccat tcacattgag agctgggtgt gtctgaagtt ttcgatatcc agtgttaacc 1620
tgtactctct cctgcgaacc tacacaccaa agctttattt atatcatccc agtatcaatg 1680
ctacacagtg ttgtcccgag cgcggggagg gcttgggcag aaacccctcg gaatgcttcc 1740
gagcagcgtg tagggatagg gaagaaccca gcaccactaa taaagctgct gcttggctgg 1800
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1860
aaaaa 1893

<210> 224
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_007274

<400> 224
acctacacac caaagcttta ttatatcat tccagtatca atgctacaca gtgtgtgcc 60

<210> 225
 <211> 4157
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007315

<400> 225

agcggggcgc	ggcgccagcg	ctgccttttc	tcccgccggg	tagtttcgct	ttcctgcgca	60
gagttctcgg	aggggctcgg	ctgcaccggg	gggatcgcgc	ctggcagacc	ccagaccgag	120
cgagagcgag	ccagcgcgct	cgggagagcg	tgccaccgcg	cgcccccgcc	tagcccttcc	180
ggatcctcgc	cgagaaaag	tttcatttgc	tgatgccat	cctcgagagc	tgcttaggtt	240
aacgttcgca	ctctgtgtat	ataacctcga	cagttctggc	acctaacgtg	ctgtgcgtag	300
ctgctccttt	ggttgaatcc	ccagggccctt	gttggggcac	aaggtggcag	gatgtctcag	360
tggtacgaac	ttcagcagct	tgactcaaaa	ttcctggagc	aggttcacca	gctttatgat	420
gacagtttct	ccatggaaat	cagacagtac	ctggcagcag	ggttagaaaa	gcaagactgg	480
gagcagctgc	ccaatgatgt	ttcatttgcc	accatccgtt	ttcatgacct	cctgtcacag	540
ctggaatgat	aatatagtcg	cttttcttgc	gagaataact	tcttgctaca	gcataacata	600
aggaagaaga	agcgtaactc	tcaggataat	tttcagggaag	acccaatcca	gatgtctatg	660
atcattttaca	gctgtctgaa	ggaagaaaag	aaaattctcg	aaaacgcccc	gagatttaaat	720
caggctcagc	cggggaatat	tcagagcaca	gtgatgttag	acaaacagaa	agagcttgac	780
agtaaatgca	gaaatgtgaa	ggacaaggtt	atgtgtatag	agcatgaaat	caagagcctg	840
gaagatttac	aagatgaata	tgacttcaaa	tgcaaaaacct	tgacagacag	agaaacagag	900
accaatggtg	tgccaaagag	tgatcagaaa	caagaacagc	tgttactcaa	gaagatgtat	960
ttaatgcttg	acaataagag	aaaggaagta	gttcacaaaa	taatagagtt	gctgaattgtc	1020
actgaactta	cccgaaatgc	cctgattaat	gatgaactag	tggaagtggaa	gcggagacag	1080
cagagcgccct	gtattggggg	gcgcgccaat	gcttgcttgg	atcagctgca	gaactggttc	1140
actatagttg	cggagagctc	gcagcaagtt	ccgcagcagc	ttaaaaagtt	ggaggaatgt	1200
gaacagaagt	acccctacga	acatgacctc	atcacaaaaa	acaaacaagt	gttatgggac	1260
cgcacacttc	gctctttcca	gcagctcatt	cagagctcgt	ttgtgggtgc	aagacagccc	1320
tgcatgccaa	cgccaccctca	gaggcccgct	gtcttgaaga	cagggggtcca	gttcaactgtg	1380
aagtttgagac	tgttggtgaa	attgcaagag	ctgaattata	atttgaagt	caaatcttta	1440
tttgataaag	atgtgaatga	gagaaatata	gtaaaaaggt	ttaggaagtt	caacattttg	1500
ggcagcgaca	caaaagtgat	gaacatggag	gagtcacaca	atggcagctc	ggcggtgaa	1560
tttcgcccac	tgcaattgaa	agaacagaaa	aatgctggca	ccagaaacgaa	tgagggtcct	1620
ctcatcggtta	ctgaagagct	tcactccctt	agttttgaaa	cccaattgtg	ccagcctggt	1680
ttgttaattg	acctcgagac	gacctctctg	cccgttgtgg	tgatctccaa	cgtagccagc	1740
ctcccgagtg	gttgggcctc	catcctttgg	tacaacatgc	tggtgvcgga	accagggaat	1800
ctgtcctctc	ctctgaactc	accatgtgca	cgatgggctc	agctttcaga	agtgctgaat	1860
tggcagcttt	ctctctgcac	caaaaagagt	ctcaatgtg	accagctgaa	ctgtgtggga	1920
gagaagcttc	ttgggtcctaa	cgccagcccc	gatggtctca	ttccgtggac	gaggttttgt	1980
aaggaaaaata	taaatgataa	aaaatttccc	ttctcgcttt	ggattgaaag	catcctagaa	2040
ctcatataaaa	aaacacctgct	ccctctctcg	aatgatgggt	gcacataggg	cttcacagc	2100
aaggagcgag	agcgctgcctt	gttgaaggac	cagcagccgg	ggaccttctc	gctgcgggtc	2160
agtgagagct	cccgggaaag	ggccatcaca	tacacatggt	tgagcggctc	ccagacagga	2220
ggcgaaacctg	acttccatgc	ggttgaaccc	tacacgaaga	aagaaccttc	tgctgttact	2280
ttccctgaca	tcatttcgaa	ttacaaaagtc	atggctgctg	agaaattccc	tgagaattccc	2340
ctgaagtata	tgtatccaaa	tattgacaaa	gacctatgct	ttggaaagta	ttactccagg	2400
ccaaaggaga	ccacagagcc	aattggaaact	gatggccctca	agggaaactg	atatatacag	2460
actgaattga	cttctgtgtc	tgaagttcac	ccttctagac	ttcagaccac	agacaaacctg	2520
ctccccatgt	ctcctgagga	gttttgacag	gtgtctcgga	tagtgggctg	tactgaattc	2580
gacagtatga	tgaacacagt	atagagcatg	aattttttct	atctctctgt	ggcagagttt	2640
tcctttcat	ctgtgattcc	ctcctgctac	tcgtttctct	cacactctgt	gtttcttagg	2700
aaatgaaaga	aaggccagca	aatctgctgc	aaactgttga	tagcaagtga	atttttctct	2760
aactcagaaa	actcagttac	ctctgaaggc	ctcatgcatc	ttactgaagg	taaaattgaa	2820
agggactatc	tgaagagtg	gtttcacaa	tgaaaaacat	ccagatacac	ccaaagttac	2880
aggaacgaaa	tgaaggtcct	gttggaaaag	agaaagtaag	caacacttag	caaatgttat	2940
gcataaagtc	agtgcccaac	tggtataggt	tggtggataa	atcagtggtt	taattggaaa	3000
ctgcttgagc	taggaacggt	aaatttctgt	ggggaatttc	ttacatgttt	tcttctgtctt	3060

aagtgttaact	ggcagttttc	cattgggttta	cctgtgaaat	agttcaaaagc	caagttttata	3120
tacaattata	tcagtcctct	ttcaaaaggtat	gccatcatgg	atctggtagg	gggaaaatgt	3180
gtatttttatt	acatctttca	cattggctat	ttaaagacaa	agacaaaattc	tgtttcttga	3240
gaagagaata	ttagctttac	tggttggtat	ggcttaatag	cactagctaa	tatacaatga	3300
aggatgtaca	tttccaaatt	cacaagttgt	gtttgatatac	caaagctgaa	tacattctgc	3360
tttcatcttg	gtcacataca	attattttta	cagttctccc	aggggagtta	ggctattcac	3420
aaccactcat	tcaaaagttg	aaattaaacca	tagatgtaga	taaactcaga	aatttaattc	3480
atgttttctta	aatgggctac	ttgttccttt	ttgttattag	ggtgggtatt	agtcatttag	3540
ccacaaaaat	gggaagggag	tagaaaaagc	agtaactgac	aacttggaata	atacaccaga	3600
gataaatga	gaatcagatc	atttcaaaac	tcatttccta	tgtaactgca	ttgagaactg	3660
catatgtttc	gctgatatac	gtgtttttca	catttgcgaa	tggttccatt	ctctctcctg	3720
tactttttcc	agacactttt	ttgagtggaat	gatgtttcgt	gaagtatact	gtatttttac	3780
cttttttctt	ccttatcact	gacacaaaaa	gtagattaa	agatgggttt	gacaagggtc	3840
ttccctttta	ctactactgc	tctatgtggc	tgatcttgt	ttttccacta	ctgctaccac	3900
aactatatta	tcactgcaat	gctgtattct	tctttggtgg	agataaaagt	ttcttgagtt	3960
ttgttttaaa	actaaagcta	aagtatctgt	attgcattaa	ataataatg	cacacagtcg	4020
tttccgtggc	actgcataca	atctggaggcc	tctctctca	gtttttatat	agatggcgag	4080
aaccttaagt	tcagttgatt	ttacaattga	aatgactaaa	aaacaaagaa	gacaacatta	4140
aaacaatatt	gttttcta	4157				

<210> 226

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007315

<400> 226

atcagatcat ttcaaaactc atttccatg taactgcatt gagaactgca tatgtttcgc 60

<210> 227

<211> 1696

<212> DNA

<213> Homo sapiens

<300>

<308> NM_009587

<400> 227

caaaggactt	cctagtggtg	gtgaaaggca	gcggtggcca	cagaggcgcc	ggagagatgg	60
cttccagcgg	ttccagcgtc	cctacacotga	gtccagctgt	ccccttttct	gggactattc	120
aaggaggtct	ccaggacgga	cttcagatca	ctgccaatgg	gaccgtttctc	agctccagtg	180
gaaccagggt	tgctgtgaac	tttcagactg	gctcagttgg	aaatgacatt	gocctccaat	240
tcaacctctg	gtttgaagat	ggagggtacg	tggtgtgcaa	cacgaggcag	aacggaaagt	300
ggggggccga	ggagaggaag	acacacatgc	ctttccagaa	ggggatgcc	tttgacctct	360
gcttccgtgt	gcagagctca	gatttcaagg	tgatggtgaa	cgggatcctc	tttgtcagtg	420
acttccacgc	cgtgcccttc	caecgtgtgg	acaccatctc	cgtcaatggc	tttgtcagtg	480
tgcttcaaat	cagcttccag	aaccoccgca	cagtcctctg	tcagcctcgc	tttccacggg	540
tgccgtcttc	ccagcctgct	tggttccca	ccaggcccgag	ggggcgccaga	caaaaaacctc	600
ccggcggtgt	gcctgcacac	cgggtcccca	ttaccagagc	agctcatocac	ccagtgccaga	660
gcgcctctgg	acagatgttc	tctactcccg	ccatcccaac	tatgatgtac	ccccaccctg	720
cttatccgat	gcttttcac	accacattc	tgggagggtg	gtaccatcc	aagtcacatc	780
tcctgtcagc	cactgtcctg	ccagtgctc	agaggttcca	catcaacctg	tgctctggga	840
accacatcgc	cttccacatg	aaccoccggt	ttgatgagaa	tgctgtggtc	gcgaacacc	900
agatcgacaa	ctctcggggg	tctgaggagc	gaagtctgcc	ccgaaaaatg	cccttcgttc	960
gtggccagag	cttctcagtg	tggaatttgt	gtgaagctca	ctgcctcaag	gtggccgtgg	1020
atggtcgagca	cctgtttgaa	tactaccatc	gctcgaggaa	cctgccacc	atcaacagac	1080
tggaagtggg	gggcgcacac	cagctgaccc	atgtgcagac	ataggcggtc	ctccggccct	1140
ggggccgggg	gctgggggtg	ggggcagtc	gggtcctctc	atcatcccca	tctcaggagc	1200
ccagccttct	caacctctgc	tgggatctgg	gctttaatgc	agaggccatg	tccttgtctg	1260

gtcctgcttc	tggctacagc	caccttgaa	cgagaagcg	agctgacggg	gattgccttc	1320
ctcagccgca	gcagcacctg	gggtccagc	tgtctggaatc	ctaccatccc	aggaggcagg	1380
cacagccagg	gagaggggag	gagtgggcag	tgaagatgaa	gcccatgctc	cagtcacctc	1440
ccatccccca	cgcagctcca	ccccagtc	aagccaccag	ctgtctgtctc	ctgtgtggag	1500
gtggcctcct	ccgcacctcc	tctctgacct	ttaacctcac	tctcaccttg	caccgtgcac	1560
caaccttca	ccctcctcgg	aaagcaggcc	tgatggcttc	ccactggcct	ccaccacctg	1620
accagagtgt	tctcttcaga	ggactggctc	ctttccacgt	gtccttaaaa	taaaagaaatg	1680
aaaatgcttg	ttggca	1696				

<210> 228

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_009587

<400> 228

cagaggactg gctcctttcc cagtgtcctt aaaataaaga aatgaaaatg ctgtgtggca 60

<210> 229

<211> 6552

<212> DNA

<213> Homo sapiens

<300>

<308> NM_012291

<400> 229

atgaggagct	tcaaaagagt	caactttggg	actctgctaa	gcagccagaa	ggaggctgaa	60
gagtgctgct	ccagcttgaa	ggagttcctg	tccaaccttc	cagctgtggtt	tcccagcagc	120
cgatctgatg	ctgagaggag	acaagcttgt	gatgccatcc	tgagggtctg	caaccagcag	180
ctgactgtcta	agctagcttg	ccctaggcat	ctggggagcc	tgctggagct	ggcagagctg	240
gcctgtgatg	gctacttagt	gtctacccca	cagcgtcctc	ccctctaact	ggaaagaaat	300
ctctttgtct	tactcgggaa	tgctgtctga	caaggaagcc	cagaggccac	actccgcctt	360
gctcagcccc	tccatgcctg	cttggtgcag	tgctctcgcg	aggctgctcc	ccaggactat	420
gaggccgtgg	ctcggggcag	cttttctctg	ctttggaagg	gggcagaagc	cctgttgtaa	480
cggcgagctg	catttgcagc	tcggctgaag	gccttgagct	tcttagtact	cttggaggat	540
gaaagtaccc	cttgtgaggt	tectcacttt	gcttctccaa	cagcctgtgc	agcggtagct	600
gccacacagc	tattttagtc	cagtgcccat	ggtctaaatg	aagcagatgc	tgatttctca	660
gatgacctgc	tctccaggca	cgtgatcaga	gccttgggtg	gtgagagagg	gagctcttct	720
gggctttctt	ctccccagag	ggccctctgc	ctcttggagc	tcaccttgga	acaactgcct	780
cgcttttctg	ggagcgccca	ccatgacaaa	gccatcagcg	cagtgaggaa	ggctcacagt	840
tacctaaagg	acaccaatct	agcccttagc	cttcagctat	gtcagctggg	gggttaagctg	900
gtgcaggttg	gggaggaagg	acctcaggca	gtggccaaag	ttctgatcaa	ggcatcagct	960
ctctgagaga	agagtatgga	ggcaccatca	ccccactctc	gggcatttga	tgagagctgc	1020
cagttcttcc	tttcaggcct	ggaaacgaggc	accaagagggc	gctatagact	tgatgccatt	1080

ctgagcctct	ttgcttttct	tggagggtac	tgctctcttc	tgacgagctg	gcgggatgat	1140
ggtgtgtgat	ggggctctct	caagcaaacg	cagtcctttc	ttcagatgta	ctttcaggga	1200
cttcaactct	acactgtggg	ggtttatgac	tttgcccaag	cgctgcagat	agttgatttg	1260
gctgacctga	cccaactagt	ggacagttgt	aaatctacgg	ttgtctggat	gctggaggcc	1320
ttagaggggc	gtcggggcca	agagctgacg	gaccacatgg	ggatgacccg	ttcttaccac	1380
agtaattttg	cctacagctt	ctatagtca	aagctctatg	ccgaggccctg	tgccatctct	1440
gtacgcgctc	gtacgacact	gggttttggtg	aagccaggca	cttatcccgga	gggtgctcct	1500
gagaagttgc	acaggtgctt	ccggctacaa	gtagagagtt	tgaagaaact	gggttaaacag	1560
gccacgggct	ccaagatggg	gattttgtgg	ctggcagccc	tgcaacctctg	taagcctgaa	1620
cacatggctg	agccagctca	tttctgggtt	cggtcaaga	tgatgcggcg	caagggtgga	1680
gacaaggagc	tacagctaaa	gactctgcga	gacacgctca	gtggctggga	ccggagagac	1740
ctggccctcc	tgctgaggga	ggagctgcag	gcatcaagg	cggtgcggcg	cgacactgga	1800
caggaacgct	tcaacatcat	ctgtgacctc	ctggagctga	gccccaggga	gacacacagc	1860

ggggcctggg	cacgagccac	ccacctggta	gaactggctc	aggtgctctg	ctaccacgac	1920
tttacgcagc	agaccaactg	ctctgctctg	gatgctatcc	gggaagccct	gcagctctctg	1980
gactctgtga	ggcctgaggg	ccaggccaga	gatcagcttc	tggaagataa	agcagagccc	2040
ttgctgtggc	tttaccatctg	tactctggaa	gccaaaatac	aggaaggtat	cgctgcgggat	2100
cggagagccc	agggccctcgg	taacttggag	gaatttgaag	tcaatgacct	gaactatgaa	2160
gataaacctcc	aggaagatcg	tttctatac	agtaacattg	ccttcaacct	gctgtcgaagt	2220
ctgctgcagt	ccaaatgcct	ggaccaagcc	ctggccctgt	gggaaggagct	gcttacaagg	2280
gggacgagccc	cagctgtacg	gtgtctccag	cagacagcag	cctcactgca	gatcctagca	2340
ggccctctacc	agctgggtggc	aaagcccatg	caggctctgg	aggtctctct	gctgctacgg	2400
attgtctctg	agagactgaa	ggaccactcg	aaggcagctg	gctctctctg	ccacataccc	2460
cagctcctcc	tgaccctcgg	ctgtccccag	tatgccagct	tacacctgga	agaggcagca	2520
tcgagcctga	agcatctcga	tcagactact	gacacatacc	tgtccttttc	ctcgactcgt	2580
gatctgcctc	gaagtcaact	ctactggact	cacocagaag	tgaccaaggg	tgtctctctg	2640
ctgctgtctg	tgccttcggga	tcttgcctcc	cagaagtctc	ccaaggcttg	gtacttgcgt	2700
cgctgtccagg	tctctgcagct	ggtggcagct	taccttagcc	tccgtcctaa	caacctctca	2760
cactccctgt	gggagcagct	ctgtgcccac	ggctggcaga	cacctgagat	agctctcata	2820
gactcccata	agctctcccg	aagcatcatc	ctactgctga	tgggcagtga	catctctctca	2880
actcagaag	cagctgtgga	gacatcgttt	ttggactatg	gtgaaaaatc	ggtacaaaaa	2940
tggcagagttc	tttccagaggt	gctgagctgt	tcagagaagc	tggcttgcca	ctctggccgc	3000
ctgggtagct	tgaagtgaagc	caaggccttt	tgtctggagg	ccctaaaaat	tacaacaaag	3060
ctgcagatac	cacgcgtgac	tgcctgttct	ctggtgctga	agggcgagct	gagctgtggc	3120
cgcaatgaca	tgtctctctg	tcagctggac	ctgcagcagg	ttctgttctt	gctttagctgt	3180
tgccagagtg	ttgagtgagg	gaactcagac	ctggactctg	tgaagaaggt	ccactcgacg	3240
aaagggaagc	agcaggtccc	ggtccctctg	cctccacagc	tcccagagga	ggagctcttc	3300
ctaagagggc	ctgctctaga	gctggtggcc	actgtggcca	agggcctgg	cccataagca	3360
ctctctacaa	actcctcccc	agttctgaaa	accagccccc	agccataacc	caacttctctg	3420
tcccattcac	ccacctgtga	ctgctcgctc	tgcgccagcc	ctgtctctac	agcagttctg	3480
ctgcgctggg	tatttggtcac	ggcaggggtg	aggtctggcca	tgggccacca	agccccaggt	3540
ctgaattctg	tgcaggtcgt	gctgaagggc	tgtctgtgaag	ccgctgagcg	cctcacccaa	3600
gctctccaag	cttccctgaa	tcataaaaaa	ccccctccat	tggttccagc	cccttgggat	3660
gagatctctg	ctcaagcata	cacactgttg	gcaactggag	gcctgaacca	gcatcaaac	3720
gagagctctg	ataaggttct	acagctcagg	ctgaagtttg	tagcagcaag	gataccccac	3780
ctagagccct	ggcgagccag	cctgtctctg	atttggggcc	tcacaaaaat	aggtggccctc	3840
agctcgtgta	ctacccaact	ttttgcaagc	tctgggggct	gcagcccaac	atttaataaaa	3900
agtgtccctg	gctcagagcc	ctctaaagct	caggggccaaa	aacgtttctg	acgagggcgc	3960
caaaaagttag	cctctgtctcc	cctgcgctcc	aataataact	ctcagaaggg	tctggaaagt	4020
agaggactgc	cctgcaacac	taaaccccca	gaccggatca	ggcaagctgg	ccctcatgtc	4080
cccttcacgg	tgtttgagga	agtcctgccc	acagagagca	agcctgaagt	acccccagcc	4140
cccgaggtac	aacagagagt	ccagacgcgc	ctcaaggtga	acttcaagtga	tgacagtgac	4200
ttggaagcac	ctgtctcagc	tgaggcctgg	ctggcagagc	agcctaaagc	acggggcact	4260
gcttcccggg	ggcggggggc	agcaagggaag	ggcctgagcc	taaagacgga	tgcgttgtt	4320
gccccaggta	gtgccccctg	gaacctggc	ctgaatggca	ggagccggag	ggccaaagag	4380
gtggcatcaa	gacattgtga	ggagcggcgt	ccccagaggg	ccagtgacca	ggccaggcct	4440
ggccctgaga	tcatgaggac	catcctctgag	gaagaaactg	ctgacaaactg	gagaaaactg	4500
agccttgaga	tctctgaggg	ctctgacggg	gaagactcag	cctcaggtgg	gaagactcca	4560
gctccggggc	ctgaggcagc	tcttggaaga	tgggagctgc	tgagcctgga	ttcagccaag	4620
aaagaagctgc	ccagcccatg	cccagacaag	gagagtgaac	aggaaccttg	tctctggctc	4680
cagctccccc	cagcccctgt	agccactggc	gttcttccac	tggactccat	ctgtgcattcc	4740
ctgagtggtt	ctttccgggg	catatgtcac	tgtctctcta	tgtggctcta	tgcccacctc	4800
tgcgccttcc	tgtccttgtg	ctggggccac	cgggatcctt	atgccactgc	ttctctgttc	4860
accagctctg	tctccatcac	ctgtgcgcc	cagctgtcca	cccactccca	cagacagctc	4920
agcacaagccc	agaagcaccc	aggatcaett	gaatatagcag	accagctgca	ggggctgagc	4980
cttcaagctg	tgcctggaga	tgtccccctg	gcccgcattc	agcgctcctt	ttctctcagg	5040
gctttggaat	ctggccactc	ccccagcct	gaaaaggaga	gtttccaggga	ggcctgtgct	5100
ctgtatcccca	gtggggtgac	tgtgtgtgtg	tgtggccctg	ccacctccca	ggccggaaac	5160
gtgggcaaca	ccctctcgtc	gaccggctgc	gaaaaggaca	gtccccagct	cagttgtcag	5220
atttccactg	ggcagaaacaa	gcttcatctg	cgttcagctc	tgaatgagtt	tgatgccatc	5280
cagaagagcac	aaaaagagaa	cagcagctgt	actgacaagc	gagaatgggt	gacagggggg	5340
ctggcactgg	accacaggat	ggaggtttctc	atcgcttccc	tagagaagtc	tgtgtgggac	5400
tgttgaagag	ggctgtcgtc	gcgctccagt	gaggagcccg	gcctgcacca	ggaggcctcc	5460

```

cgccctacagg agctgctaca ggactgtggc tggaaatcac ctgaccgcac tctgctgaaa 5520
atcatgctca gtgtgtccgg tgccctcacc cctcaggaca ttcaggccct ggccctacggg 5580
ctgtgcccaa ccagcccaaga gcgagcccaag gagctcctga atgaggcagt aggcagctcta 5640
caggcgctga cagtaccaaag caatagccac cttgtcttgg tcttagacaa ggactgtcag 5700
aaagtgcogt gggaaagcat gccagccctc caagcactgc ctgtcaccgg gctgcctccc 5760
ttccgtctcc tactcagcta ctccatcctc aaagagtatg gggcctcgcc agtgcgtgagt 5820
caagggtgtg atccacgaag tacctcttat gtcctgaacc ctcaacaata cctgtcaagc 5880
acagaggagc aatttcgagc caatttcagc agtgaagctg gctggagagg agtggttggg 5940
gaggtgccaa gacctgaaca ggtgcaggaa gccctgacaa agcatgattt gtatatctat 6000
gcagggcagt gggctgtgtg ccgcttccct gatgggcagg ctgtcctgcg gctgagctgt 6060
cgggcagtg cctcgtctgt tggctgttag agtgccggcc tggctgtgca tggaaacctg 6120
gagggggctg gcctcgtgct caagtacatc atgctcgtgt gccctctgtt tctgggtaat 6180
ctctgggagt tgactgaccg cgacattgac cgctacacgg aagctctgct gcaaggctgt 6240
cttggagcag gccagggggc ccccttctc tactatgtaa accaggcccg ccaagctccc 6300
cgactcaagt atcttattgg ggcctgcacc atagcctatg gcttgccctg ctctctcggg 6360
taaccctcat gagctgtctt attgatgcta gaagcctcat aactgttcta cctccaaggc 6420
tagatttaac ccttaggata actcttttaa agtgattttc cccagtgttt tatatgaaac 6480
atttcccttt gatttaacct cagtataata aagatacac attttaaacc tgaaaaaaaa 6540
aaaaaaaaaa aa 6552

```

```

<210> 230
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_012291

```

```

<400> 230
agcctcataa ctgttctacc tccaaggtta gatttaatcc ttaggataac tcttttaagg 60

```

```

<210> 231
<211> 6317
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_013261

```

```

<400> 231
tagtaagaca ggtgccttca gttcactctc agtaaggggc tgggtgcctg catgagtgtg 60
tgctctgtgt cactctggat tggagttgaa aaagcttgac tggcgtcatt caggagctgg 120
atggcgtggg acatgtgcaa ccaggactct gagtctgtat ggagtgaact cgagtgtgct 180
gcctcgtgtg gtgaagacca gcctctttgc ccagatcttc ctgaacttga tctttctgaa 240
ctagatgtga acgacttggg tacagacagc tttctgggtg gactcaagtg gtgcagtgac 300
caatcagaaa taatatccaa tcagataaac aatgagcctt caaacatatt tgaaagata 360
gatgaagaga atgaggcaaa ctgtctagca gtccctcacg agacactaga cagtctccct 420
gtggatgaag accggtgtcc ctcatgtgat gcgctgacag atggagacgt gaccactgac 480
aatcaggctca gtccctcttc catgcctgac ggcacccttc caccaccagg ggccagaagag 540
ccgtctctac ttaagaagct cttaactggca ccagcccaaca ctcagctaaag ttataatgaa 600
tgcagtggtc tcagtaacca gaacctgca aatcacatc acaggatcag aacaaacctt 660
gcaattgtta agactgagaa ttcaatggag aataaagcga agagtatttg tcaacagcaa 720
aagccacaaa gacgtccctg ctccggagctt ctcaaatatc tgaccacaaa cgatgacctt 780
ctccacacca aaccacaga gaacagaaac agcacgacag acaaatgcac ctccaaaagg 840

```

```

aagtcaccaca cacagtgcga gtcacaacac ttacaaggca aaccaacaaa tttatctctt 900
cctctgaccc cagagtccac aaatgacccc aagggttccc catttgagaa caagactatt 960
gaacgcacct taagtgtgga actctctgga actgcaggcc caactccacc caccactctt 1020
cctcataaag ccaaccaga taacctttt agggcttctc caaagctgaa gtccctctgc 1080
aagactgtgg tgccaccacc atcaagaag cccaggtaca gtgagctctc tggtagacaa 1140

```

ggcaataact	ccaccaagaa	agggcggag	caatccgagt	tgtatgcaca	actcagcaag	1200
tcctcagtc	tcactggttg	acacgagaa	aggaagacca	agcgccccag	tcctcggtgt	1260
ttgtgtgacc	atgactattg	ccagtcgaatt	aattccaaaa	cagaataact	catttaataa	1320
tcacaggagc	tcacaagactc	tagacaacta	gaaaataaag	atgtctcttc	tgatggcgag	1380
ggcgagatt	gtttctccac	agattccagac	cagtgctacc	tgagagagac	tttggaggca	1440
agcaagcagg	ttctctcttg	cagcacaga	aaacagctcc	aagaccagga	aatcccgagcc	1500
gagctgaaca	agcaactcgg	tcattccagct	caagctgttt	ttgacgacga	agcagcaag	1560
accgtgtgaa	tgaggagcag	tgatttcagt	aatgaacaat	ttccaaaact	acctatgttt	1620
ataaattcag	gactagccat	ggatggcctg	tttgatgaca	gcgaagatga	aagtataaaa	1680
ctgagctacc	cttgggtagg	cacgcaatcc	tattcaattgt	tcaattgttc	tccttcttgt	1740
ttctctttta	actctccatg	tagagattct	gtgtcaccac	ccaaatcctt	attttctcaa	1800
agaccccaaa	ggatgcgctc	tcgtttcaagg	tcctttttctc	gacacaggtc	gtgttccoga	1860
tcaccatatt	ccaggtcaag	atcaaggctc	ccaggcgatga	gatctcttc	aagatcctgc	1920
tattactatg	agtcaagcca	ctacagacac	cgacgcgacc	gaaattctcc	cttgtatgtg	1980
agatcacgtt	caagatcgcc	ctacagccgt	cgccccaggt	atgacagcta	cgaggaaat	2040
cagcacgaga	ggctgaagag	ggaagaaat	cgacagagat	atgagaagcg	agagctctgag	2100
agggcccaag	aaaggagag	gcagaggcag	aaggcaattg	aaagcgccgc	tgtgatttat	2160
tcgtgtaaaa	tcagacgtga	cacaaacagg	tcagaactga	gggaccgttt	tgaagtgttt	2220
gggtgaattt	aggagtgcac	agtaaatctg	cggtgatgatg	gagacagcta	tggtttcaat	2280
acctaccgtt	atacctgtga	tgcttttgct	gctcttgaaa	atggatacac	tttgccgagg	2340
tcacaacgaa	ctgactttga	gctgtacttt	tggtgacgca	agcaattttt	caagtctaac	2400
tatgcagacc	tagatctaaa	ctcagatgac	ttgaccctgt	cttccacca	ggagcagtat	2460
gactctctgg	attttgatag	tttactgaaa	gaagctcaga	gaagcttcgc	caggttaaat	2520
gttccctagc	tcaggatgac	agagggatgg	cgaataacctc	atggggacgc	cggtctcttc	2580
ctaaagacta	ttgcagctoa	tacttaggaa	tttctctcac	tttactactc	ctgtacaaa	2640
acaaaacaaa	acaacaacaa	tacaacaaga	acaacaacaa	caataaaca	aatgttttac	2700
atgaacacag	ctgctgaaga	ggcaagagac	agaatgatat	ccagtaagca	catgtttatt	2760
catgggtgtc	agctttgctt	ttcctggagt	ctcttggtgc	tggagtgtgc	gtgtgtgcat	2820
gtatgtgtgt	gtgtatgtat	gtgtgtggtg	tgtgtgctgt	gtttagggga	agtatgtgtg	2880
ggtaactgtg	aggactggg	gcactgacc	agaaatgtgc	agggcaaaac	atttcgaatt	2940
gcagcagttc	catgaagaca	cgcctaaaac	ctagaacttc	aaaaatgttc	tattctatct	3000
aaaaagaaat	atatatatat	atatatatat	atatatatat	atatataaat	taaaaaggaa	3060
agaaaactaa	caaccaacaa	accaaccaac	caaccaacaa	ccaccctaaa	atgacagccg	3120
ctgatgtctc	ggcatcagcc	tttgtactct	gtttttttaa	gaaagtgcag	aatcaacctg	3180
aagcaagcct	tcctctctaa	cgtaatgatt	atgatgaact	cctgaagaaa	ccacaggttc	3240
catgaactaa	atatctctgc	tcctctctct	tcctctctct	tcctcttttt	ttttcttttt	3300
ctttttgcca	tggaaactgg	gtgggagagg	atactgcggg	caccagaatg	ctaaaagttc	3360
ctaaactttt	gaagtttctg	tagtttcaatc	ttaatcctga	ccaccatgta	aatgtccaaa	3420
atgtgtgact	tcactgcgaa	atttcaaaag	cgctgtcaat	ggtaacagcg	cagactgtgt	3480
cagcggttct	ttctgagcag	cggaacccgg	gttacattac	taatgagagt	tggtgtgaac	3540
tcctcagagt	gtgttcagat	agtgtaattg	ctacattctc	tgatgtagt	aagtattttac	3600
agatgttaaa	tggagatttt	ttatttttatg	tataataact	acaacaatg	ttttttttgt	3660
tcacgtctatg	cactgttaaat	gcagccttct	tttcaaaact	ctcaaaattt	tcattatcaa	3720
gaatttctaa	atgtaattat	gagggtgaac	aattattgta	caactaacata	tttagaagct	3780
gaacttactg	cttatataata	tttgattgta	aaaaacaaaa	gacagtgtgt	gtgtgtgtgt	3840
agtgcaacaa	gagcaaaatg	atgtcttccg	cacatccatc	ccttaggtga	gcttcaatct	3900
aagcatcttg	tcagaataata	tcctagtccc	ctaaaggatt	taaccacttc	tgcgatattt	3960
ttccacattt	tccttgctgc	tgttttttct	tgaagtttta	tacactggat	ttgttagggg	4020
aatgaatttt	tcctactctaa	aattttttct	gaagatatca	tgattttatg	taaggtctct	4080
caatgggttaa	ccattgaagaa	atgtttttat	tttctctatc	cacactgtag	ttgaaactag	4140
aagtcaaaaa	ctttttttaa	atgctgtttt	gttttaattt	ttgtgatttt	aatttgatac	4200
aaaaatgctga	ggttaataatt	atagtatgat	ttttacaata	attaatgtgt	gctggaagac	4260
tatctttgaa	gccaagattt	ctttccctgt	gcagagtatg	acgatggat	ttatctgtat	4320
tttttacagt	tatgcactct	gtataaaatc	tgactatttca	ttcttttgtt	tactaaagag	4380
acatatattat	cagttgcaga	tagcctattt	attataaatt	atgagatgat	gaaaataata	4440
aagccaggtg	aaattttctca	octaggatgc	atgacaattg	tcagggttga	gtgtgaagtc	4500
ttcattttggg	aaattcagct	tttgcaagag	cagtggtttc	acttgcaata	gcaatggctc	4560
tgactgtgacc	atgggtgtgt	tcctgatgac	attgctctct	ctaaatttaa	taaaaacttc	4620
agaaaaacct	ccattttgat	catcaggatt	tcactgtagt	gtggagtccc	tggaattgaa	4680
ttcagtaaca	tttgaggtgt	gtattcaagt	ttctaaattg	agattcgatt	actgtttggc	4740


```

tgacatgact tttctggaag acatgatata cctactactc aattgttctt ttcoctttctc 4800
tcgcccaca cgatcttgta agatggattt cacccccagg ccaatgcagc taattttgat 4860
agctgcattc attatcacc agcatattgt gttctgagtg atccactgt ttgtcctgtc 4920
ggatgcttgc ttgatttttt ggtctcttat ttctaagtag atagaaagca ataaaaatcac 4980
tatgaaatga aagaacttgt tcacaggttc tgcgttacaa cagtaacaca tctttaatcc 5040
gcctaattct tgttgttctg taggttaaat gcaggtattt taactgtgtg aacgcaccaac 5100
taaaagttaac agtctttctt tctgaatttt agatctctc ttgtgtagaa taataataaaa 5160
aagactatca agagcaataa attattttta aaaaactcag atttagtaaa tctctattatg 5220
tgttcaagga ccaactgtgt tctctatttt gcctttaaatt ttttgtgaac caatttttaa 5280
tacattctcc tttttgcctt ggattgttga catgagtggg atacttggtt tcttttctta 5340
cttatcaaaa gagacaccta cagatatcat attgaggatt aatttatccc ccttaccccc 5400
agcctgacaa atattgttac catgaagata gttttcctca atggacttca aattgcactt 5460
agaattagtg gagcttttgt atctctcgca gacactgtgg gttagcccatc aaaatgtaag 5520
ctgtgctctc ctcaatttta tttttatttt ttggggagag aattttcctc atgaacacgt 5580
gcaccccatc atcaactggag gcaaatttca gcatagatct gtaggatttt tagaagaccg 5640
tgggcatgtg ccttcatgcc ttggttaagta ccacatctac aatttttgta accgaactgg 5700
tgctttagta atgtgtgatt ttttctttt taaaagagat gtacgagaat aattcttcca 5760
gtgcaacaaa atcaattttt tgctaaacga ctccgagaac aacagtgggg ctgtcaacat 5820
tcaaaagcagc agagagggaa ctttgaccta ttgattgtga gcttccagta ggtaatgatt atgtgtcctt 5880
aaggaaacct tttcatgcct ttagatgtga gcttccagta ggtaatgatt atgtgtcctt 5940
tcttgatggc tgtaatgaga acttcaatca ctgtagtcta agacctgac tatagatgac 6000
ctagaatagc catgtactat aatgtgatga tctctaaatt gtacctatgt gacagacatt 6060
ttcaataatg tgaactgctg atttgatgga gctactttaa ttgttgtag tgaaagtgtg 6120
atactgttgg ttgaactatg ctgaagaggg aaagtgcagg attagttag ccctgcggcg 6180
gccctttttc caccctgcaa ttctacatgt attgtgtgtg tttatttcat tgtatgaaa 6240
ttcctgtgat tttttttaa ttgtgcagtc acatcagcct cactgagcta ataaagggaa 6300
acgaatgttt caaatct 6317

```

<210> 232

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013261

<400> 232

ctgtagtcta agacctgac tatagatacc tagaatagcc atgtactata atgtgatgat 60

<210> 233

<211> 3237

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013277

<400> 233

```

gcgaagtga ggtggggcca ggtggggcca ggtgactga atgtatctcc tagctatgga 60
ctaaataata catgggggga aataaacagc tatctcatgag ggtgaaatg tgacccagca 120
ggaataattac aactattttc aattgacgtt gaataggatg agtcatggaa tttaatgtat 180
ttactgaaga ttatactact ggtagataga agagctaaag aaagtatgat actatgatgc 240
tgaattgtcg gaatctgttt gagcagcttg tgcgcgggt ggagattctc agtgaaggaa 300
atgaagtcca attatccag ttggcgaaag actttgagga tttccgtaaa aagtggcaga 360
ggactgacca tgaactgggg aaatacaagg atcttttgat gaaagcagag actgagcgaa 420
gtgctctgga tgttaagctg aagcatgcac gtaatcaggt ggatgtagag atcaaacgga 480
gacagagagc tgaggctgac tgcgaaaagc tggaacgaca gattcagctg atttcgagaga 540
tgctcatgtg tgacacatct ggcagcattc aactaacgca ggagcaaaaa tcagctctgg 600
ctttctcaa cagaggccaa ccatccaaga gcaatgctgg gaacaaaaga ctatcaacca 660
ttgatgaatc tggttccatt ttatcagata tcaagcttga caagatgatg gaactactgg 720
attgggactc tcttttggg aagactttca aactgaagaa gagagaaaaa aggcgctcta 780

```

ctagccgaca	gtttgttgat	ggteccctcg	gacctgtaaa	gaaaactcgt	tccattggct	840
ctgcagtaga	ccagggggaat	gaatccatag	ttgcaaaaac	tacagtgact	gttcccaatg	900
atggcgggcc	catcgaaagt	gtgtccacta	ttgagactgt	gccatattgg	accaggagcc	960
gaagggaaaac	aggtacttta	caaccttgga	acagtgactc	cacctgtaac	agcaggcagc	1020
tggaagcaag	aactgagaca	gacagtgtgg	gcacgcacac	gagtaaatgga	gggatgcccc	1080
tgcatgactt	tggtttcaag	acggttatta	aaacctgaatc	ctgtgttcca	tgtgtgaaagc	1140
ggataaaatt	tgggaaaatta	tctctgaagt	gtcgagactg	tcgtgtgtgc	tctcatccag	1200

aatgtcgagg	ccgctgtccc	cttccctgca	ttcctaccct	gataggaaac	cctgtcaaga	1260
------------	------------	------------	------------	------------	------------	------

ttggagaggg	aatgtctgca	gactttgtgt	cccagacttc	tccaatgac	ccctccattg	1320
ttgtgcatgt	tgtaaatgag	attgagcaaa	gaggtctgac	tgagacaggc	ctgtatagga	1380
tctctgtgct	tgaccgcaca	gtaaaagagc	tgaaaagaaa	attcctcaga	gtgaaaactg	1440
taccctcctc	cagccaaagt	gatgatattc	atgctatctg	tagcctctta	aaagactttc	1500
ttcgaaacct	caaaagaaact	cttctgacct	ttcgccctaa	cagagccttt	atggaagcag	1560
cagaatacac	agatgaagac	aacagcatag	ctgccaatga	ccaagctgtt	ggtgaactgc	1620
cccaggccaa	cagggacaca	ttagctttcc	tcatgattca	cttgacagag	gtggctcaga	1680
gtccacatac	taaaattggat	gttgccaatc	tggttaaagt	ctttggccct	acaatagtgg	1740
cccatgctgt	gcccaatcca	gcgccagtga	caatgttaca	ggacatcaag	cgtcaaccac	1800
aggtggttga	gcgcctgctt	tccctgcctc	tgagattatg	gagtcagttc	atgatggtgg	1860
agcaagagaa	catlgacccc	ctacatgtca	ttgaaaactc	aaatgccttt	tcaacaccac	1920
agacaccaga	tattaaagt	agtttactgg	gacctgtgac	cactcctgaa	catcagcttc	1980
tcaagactcc	ttcatctagt	tccctgtcac	agagagtcgg	ttccacccct	accaagaaca	2040
ctctatgatt	tgaggagcaaa	agcaagtcct	ccactaacct	aggacgacaa	ggcaactttt	2100
ttgcttctcc	aatgctcaag	tgaaagtcac	ttgcctgtgt	acttcccgac	atttgactgac	2160
tataaagaag	gacacatctg	tactctgctc	tgacgctccc	tgtaactcatt	actactttta	2220
gcatttctcca	ggcttttact	caagtttaat	tgtgcatgag	ggttttatta	aaactatata	2280
tatctccctc	tccttctcct	caagtcacat	aatatcagca	ctttgtgctg	gtcattgttg	2340
ggagcttttta	gatgagacct	ctttccaggg	gtagaagggt	tagtatggaa	ttggttgtga	2400
tctctttttg	ggaagggggt	tattgttctc	ttggcttaaa	gccaaatgct	gctcatagaa	2460
tgatctttct	ctagtttcat	ttagaactga	tttccgtgag	acaatgacag	aaacccctacc	2520
tatctgataa	gattagcttg	tctcagggtg	ggagatggga	gggcagggca	aagaagggat	2580
tagaccagag	gttttaggat	gcctccttct	aagaaccaga	agttctcatt	ccccattatg	2640
aactgagcta	taatatggag	ctttcataaa	aatgggagtc	attgaggaca	gaactagtga	2700
tgaggagtatg	cgtagctttg	atttgtagta	ttaggtcttt	aatagtgttg	agtggcacaa	2760
ccttgtaaat	gtgaaagtac	aactcgtatt	tatctctgat	gtgccgctgg	ctgaactttg	2820
gggtcatttg	gggtcaaaagc	cagtttttct	tttaaaattg	aattcattct	gatgcttggc	2880
ccccatcccc	ccaaccttgt	ccagtgagagc	ccaacttcta	aaggtcaata	tatcatcctt	2940
tgcatcccca	actaacaata	aagagtgggc	tataagggaa	gattgtcaat	attttgttgt	3000
aagaaaagct	acagctcatct	tttctttgca	ctttggatgc	tgaaaatttt	cccatgggaa	3060
atagccacat	ctagatagat	gtgagctttt	tcttctgtta	aaattattct	taatgtctgt	3120
aaaaacgatt	ttcttctgta	gaatgtttga	cttcgtattg	acccttatct	gttaaacacc	3180
tatttgggat	aatatttggg	aaaaaagtaa	atagcttttt	caaaatgaaa	aaaaaaa	3237

<210> 234

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013277

<400> 234

ctcattcccc attatgaact gagctataat atggagcttt cataaaaatg ggatgcattg 60

<210> 235

<211> 1122

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013409

<400> 235

```

gctcctcgcg ccgcgcctgc cccaggatg gtccgcgcga ggcaccagcc ggggtgggctt 60
tgctcctcgc tgctgtcgtc ctgcagttc atggaggacc gcagtgccca ggtcgggaac 120
tgctggctcc gtcaagcgaa gaacggccgc tgccaggtcc tgtacaagac cgaactgagc 180
aaggaggagt gctgcagcac cgccggcgtg agcacctcgt ggaccgagga ggacgtgaat 240
gacaaacacac tcttcaagtg gatgattttc aacggggcgc cccccaactg catccctgt 300
aaagaaacgt gtggaacgt ggactgtgga cctgggaaaa aatgccgaat gaacaagaag 360
aacaaccccc gctgcgtctg cgccccgat tgttccaa caacctggaa ggggtccagtc 420
tgctgtccag atgggaaaaa ctaccgcaat gaatgtgcac tctaaggc aagatgtaaa 480
gagcagccag aactggaagt ccagtaccaa ggcagatgta aaaagacttg tcgggatgtt 540
ttctgtccag gcaagctccac atgtgtggtg gaccagacca ataatgccta ctgtgtgacc 600
tgtaatcgga ttgcccgaga gcctgccttc tctgagcaat atctctgtgg gaatgtgga 660
gtcacctact ccagtcgctg ccacctgaga aaggctacct gcctgctggg cagatctatt 720
ggattagcct atgagggaat gtgtatcaaa gcaaaagtcct gtgaagatat ccagtgcact 780
gggtggaaaa aatgtttatg ggatttcaag gttgggagag gcggtgtctc cctctgtgat 840
gagctgtgcc ctgcagtaaa gtctgtctgt ccagtgcact tgccacttat 900
ccagcggagt gtgccatgaa ggaagctgcc tgctctcag gtgtgctact ggaagtaaa 960
cactccggat cttgcaactc catttcggaa gacaccgagg aagaggagga agatgaagac 1020
caggactaca gttctctat atctctatt cttagtggtt aaactctcta taagtgttca 1080
gtgttcacat agcctttgtg caaaaaaaaa aaaaaaaaaa aa 1122

```

<210> 236

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013409

<400> 236

```

gaagatgaag accaggacta cagctttcct atatcttcta tcttagagtg gtaaaacttc 60

```

<210> 237

<211> 11389

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014246

<400> 237

```

atggcgccgc cgcgcgcgcg cgtgctgccc gtgctgtgcg tctggccgcg cgccgcgcgc 60
ctgcgcgcga tgggggtgcg agcggcgcgc tgggagccgc gcgtaccgcg cgggaccgcg 120
gccttgcgcc tccggcccg gctgtacctac gcgggtggcg ccgcttgacc gccccggggcg 180
ccgcgggaag tgctggacgt gggccgcgat gggcggtcgg caggacgtgc gcgcgtctcg 240
ggcgcggggc gcgcgctgcc gctgcgaagt cgccttggtg ccgcagctgc cccgacggcg 300
ctgagccgcc gcctgcggcg gcgcacgcac ctcccggtg ccggagcccg tgcccgcctc 360
tgccgaaccc gtgcggcgct ctgcggggcg cctctgtccc ccgtccccgc cagctgcgcg 420
gcgcgcgacg attcggcgct cgcagctccg accaccttac ccgcctgcgc ctgcgcgcgc 480
cgccccaggg cccgctgtcc cggccgctccc atctgcctgc cgccggcgcg cctcgtccgc 540
ctgcgtctgc tgtgcgccct gcgcgcgcgc gctggcgccg tccgggtggg actggcgctg 600
gagcgcgcga agtctccgat gcctccgcgc tgcacctccc catcgccgcc cctgccgcgc 660
aaacttgcgc aagccccggc gggccggcg cgacgggccc ggccgggacg gagcggcaga 720
gggagcctga agtttccgat gcccaactac caggtggcgt tgtttgagaa cgaacccgcg 780
ggcaccctca tcttccagct gcacgcgcac tacaccatcg agggcgagga ggagcgcgtg 840
agctattaca tggaggggct gttgcagcag cgctcccggg gctacttccc aatcgactct 900
gccacggcgc ccgtgagcac ggacagcgta ctggaccgcg agaccaagga gacgcacgtc 960
ctcagggtga aagccgtgga ctacagtaag ccgcgcgcgt cggccaccac ctacatcact 1020

```

gtcttggcca	aagacaccaa	cgaccacagc	ccggtctctg	agcagtcgga	gtaccgcgag	1080
cgcgtgcggg	agaaacctga	ggtggggtac	gaggtgctga	ccatccgcgc	cagcgaccgc	1140
cgctgcacca	tcaacgcgcaa	cttgcgttac	cgctgttctg	ggggcgcgct	gagcgtcttc	1200
cagctcaacg	agagctctcg	cgtggtgagc	acacggggcg	tgctggaccg	ggagcagggc	1260
ccgagtgacc	agctcctggt	ggaggccaac	gaccaggggc	gcaatccggg	cccgctcagt	1320
gccacggcca	ccgtgtacat	cgaggtggag	gacgagaacg	acaactaccg	ccagttcagc	1380
gagcagaact	acgtggtcca	ggtgcocgag	gagctggggc	tcaacacggc	tgtctgcga	1440
gtgcagggca	cggaccggga	ccaggggcag	aacgcggcca	ttcaactacg	ctactctcag	1500
gggaacctgg	ccggccagtt	ctacctgcac	tgcgtgagcg	ggatcctgga	gtatctcaac	1560
cccttggatt	tcgaggaagt	ccagaataat	tgcgtgagca	ttaaaggcca	ggtatggggc	1620
cgccccccgc	tcatacaatt	ttcagggggtg	gtgtctgtgc	aggtgtctga	tgtcaaacgac	1680
aacgagccta	tctttgtgag	cagcccccttc	caggccaacg	tgtctggaga	tgtgccccctg	1740
ggctaccccc	tgtgtgcacat	tcaggcggtg	gacgcggact	ctggagagaa	cgccccgctg	1800
catcatcgcc	tgtgtggcac	ggcctccacc	tttctggggg	cgcgacgcgc	tgaccctcag	1860
aatcctgcgcc	ccaccctctga	cttccctcttc	cagatccaca	acagctccgg	ttggatccaa	1920
gtgtgtgcgc	agctggaccg	cgaggaggtg	gagcactaca	gcttcggggt	ggagcgctgtg	1980
gaccaaggct	cgccccccat	gagctcctcc	accagcgtgt	ccaacacggt	gctggacgtg	2040
aatgacaacg	accocggtgt	cacgcagccc	acctacgagc	ttcgtctgaa	tggagatgcg	2100
ggcgtggggg	cgagcgttgt	gaccctgcag	ggccgcgacc	gtgacgccaa	caagtgtgatt	2160
acctaccagc	acccatcagg	caacaccggg	aacgcgtttg	cactcagctg	ccagaggggg	2220
ggcgggccta	tcaacctggc	gctacctctg	gactacaagc	aggagcagca	gtactgtctg	2280
gcggtgacag	catccgacgg	cacacggctg	caactcgcgc	atgtcctaag	caacgctcact	2340
gtgtgcacaa	ccccacggcc	tgtctttcag	agctcccatc	acacagttag	tgtcagttag	2400
gacaggccctg	tggggcaactc	caattgtacc	ctcagtgcca	acgatgagga	caacaggagag	2460
aatgcgcctg	tcaactcagt	gattcaggag	cccgctgcgc	agtctccgat	tgaccocctag	2520
agtggaccca	tgtacaacct	gatggagctg	gactaatgga	accaggtcgc	ctacacgctg	2580
acctcatggt	ccaggagcaa	cggcatcccg	acagaatcgc	acacaccacc	ctcagagatc	2640
ctcatcctcg	atgccaatga	caatgcaccc	cagttcctgt	gggattttct	ccagcgtttcc	2700
ttctttgagg	atgctccacc	ctcgaccagg	atctctgccac	tctctgccac	ggaccggggac	2760
ttcagctccca	atggggcgtct	gctgtacacc	ttccagggtg	gggacgacgg	cgaatggggag	2820
ttacagatgc	agcccaagtc	cgggtgtgatt	cgccaccacg	ggcgctgga	ccgggagaaat	2880
gtggcgctgt	acaaaccttg	ggctctggct	gtggatcggg	gcaatccccc	ttccctttagc	2940
gctcgtgcgt	aaatccagggt	gacctatttg	gacattaatg	acaatggccc	catgttttgag	3000
aaggacgaac	tggagcttgt	tgttgaggag	aaacaaccag	tggggtcggt	ggtggccaaag	3060
atttgtgtcta	acgaccctga	tgaaggccct	aatgcccaga	tcaatgtatca	gattgtggaa	3120
ggggacatcg	ggcatttctt	ccagctggac	ctgctcaacg	gggacctgcg	tgcctatggtg	3180
gagctggact	ttgagggtccg	cggggagtat	gtgctggtgg	tgacggccac	gtcggctccg	3240
ctgggtgagcc	gagccacgggt	gcacatcctt	ctcgtggacc	agaatgacaa	cccgcctgtg	3300
ctgcgccact	tcagagctcat	cttcaacaac	tatgtcacca	acaagtccaa	caagtttcccc	3360
acggcgctga	tcgggtcctc	cccgcccatc	gaaccocgag	tgtcagacag	cttcaactacc	3420
acctctgtgc	aggggcaacga	gctgcgcctg	ttgctgtctg	accccgccac	gggcgaactc	3480
cagctcagcc	gcgacactgga	caacaacccg	cgcgtggagg	cgctcatgga	gtgtctctgtg	3540
ttctgatgcca	ttccacagcgt	cacggccttc	tgcacctcgc	gtgtcaccat	catcacggac	3600
gacatctgtga	ccaacagcat	cactgtccgc	tgaggaacaa	tgatccagga	gaagtctcctg	3660
tcoccgctgc	tggccctctt	cgtggagggg	gtggccgcgc	tgtctccacc	caaccaaggac	3720
gacgtctctgc	tctcaacgt	ccagaacgac	accagctca	gctccaacat	ctgtgaactg	3780
acctctccgc	cgtcgtctgc	tggcggcgtc	cgcgccacgt	tcttcccgta	ggaggaacctg	3840
caggagcagga	tctacctgaa	tgggacgctg	ctgaccacca	tctccacgca	cgccgtgctg	3900
gctctcgacg	acaaactctg	cctgcgcgag	ccctgcgaga	actaatgaa	gtgcgtgtcc	3960
gttctgcgat	tcgacacgct	cgcgcctctc	ctcagctcca	ccacgctgca	cttccgggcc	4020
atccacccca	tcaacggcct	gcgctgcgcg	tgcccgccgc	gcttcaaccg	cgactactgc	4080
gagacgggga	tcgaacctctg	ctactccgac	ccgtgcggcg	ccaaacggcg	ctcgcgcagc	4140
cgcgagggcg	gctcaacctg	cgaagtgttc	gaggacttca	ctggagagca	ctgtgaggtg	4200
gatgcccgct	caggcccgctg	tgccaaacggg	gtgtgcaaga	acggggggcac	ctcgtgtaac	4260
ctgctcatcg	cgggcttcca	ctgcgtgtgt	ccctcctggc	agtatgagag	ggcctactgt	4320
gaggttgacca	ccaggagcctt	cccgccccag	tcttctgtca	ccttccgggg	cctgagacag	4380
cgccttcaact	tcaacctctt	cctcacgttt	ggcactcagg	aaaggaaacg	cttgccttctc	4440
tacacgcggcc	gcttccaatga	gaagcacgac	ttcatcgccc	tggagatcgt	ggacgcagac	4500
gtcagctcca	ccttctcttgc	agggcagaca	acaacacccg	tggcaccgga	gtgtcccgat	4560
ggtgtgagtg	acgggcgggtg	gcactctctg	cagggtcagt	actcaacaaa	gcccaatat	4620
ggccacctgg	gctgcgccca	tgggcgctcc	ggggaaaaga	tggccgtggt	gacagtggat	4680

gattgtgaca	caaccatggc	tgtgcgcttt	ggaaaggaca	tcgggaacta	cagctgcgct	4740
gcccaggcca	ctcagaccgg	ctccaagaag	tcctctggatc	tgaccggccc	tcctactcctg	4800
gggggtgtcc	ccaactctgc	agaagacttc	ccagtgacaca	accggcagtt	ctggggctgc	4860
atgcgggaac	tgtcagtcga	cggcaaaaat	gtggacatgg	ccggattcat	cgccaaacat	4920
ggcaaccggg	aaggctgcgc	tgtctcggag	aaetctctcg	atgggaggcg	gtgtcagaat	4980
ggaggcagct	gtgtcaacag	gtggaaatag	tatctgtgtg	agtgtccact	ccgattcggc	5040
gtgaagaaat	gtgagcaac	catgcctcac	ccccagctct	tcagcgggtg	gagcgtcgtg	5100
tcctgaggtg	acctgaacat	catcatctct	gtgcctctgt	acotggggct	catgttcocg	5160
accggcgagg	aggacacggt	tctgatggag	gccaccagtg	gtggggccac	cagctttcgc	5220
ctccagatgc	tgaacaacta	ctccagtttc	gaggtgtccc	acggcccttc	cgatgtggag	5280
tcctgatgatc	tgtccgggtt	gggggtgacc	gacggggagt	ggcaccacct	gctgatcgag	5340
ctgaagaatg	ttaaggagga	cagtgaagat	aagcacctgg	tcacatgac	cttgactat	5400
gggatggacc	agaacaaggc	agatatcggg	ggcatgtctc	ccgggctgac	ctgaaggagc	5460
gtggtggtcg	gagggcgctc	tgaagacaag	gtctccgtgc	gcctgtgatt	ccgagctgc	5520
atgcagggag	tgaggatggg	ggggacgccc	accaacgtcg	ccacctgaa	catgaacaac	5580
gcactcaagg	tcagggtgaa	ggacggctgt	gatgtggacg	acccctgtac	ctcgagcccc	5640
tgtcccccca	atagccgctg	ccacgacgcc	tgggaggact	cagctgcgt	ctgtgacaaa	5700
gggttaacttg	gaataaactg	tgtggatgcc	tgtcacctga	acccctgcga	gaactctggg	5760
gcctgcgtgc	gctcccccg	ctcccccgag	ggctacgtgt	gcgagtggtg	gcccgctcac	5820
tcaggggcgt	actgtgagaa	caaaactcgac	cttccgtgcc	ccagaggctg	gtgggggaac	5880
cccgctctgtg	gacctcgcca	ctgtgcctgc	agcaaaagct	ttgatccoga	ctgtataaag	5940
accaaagccc	agtgccaatg	caaggagaat	tactacaagc	tcctagcccc	ggacaacctg	6000
ctgcctcgct	actgcttccc	ccatggctcc	cacagccgca	cttgcgacat	ggccaccggg	6060
cagtggtgct	gcaagcccg	cgctcatcg	cgccagtgca	accgctgcga	caaccctgtt	6120
gcggaggatca	ccaacgtcgg	ctgtgaagtg	atctacaatg	gctgtcccaa	agcatttgag	6180
gcggcgatct	ggtggccaca	gaccaagtct	gggcagccgg	ctgcggtgcc	atgccttaag	6240
ggatccgttg	gaattgctgt	ccgacactgc	agcggggaga	agggctgtgt	gcggccagag	6300
ctctttaact	gtaccacat	ctctcttgtg	gacctcagg	ccatgaatga	gaagctgagc	6360
ccgaatcga	cggcgtgga	cgccgccagg	gcctgcagc	tggtgaggcg	ctgtgcagct	6420
gctacacagc	acaaggccac	gctctttggc	aatgacgtgc	gcacggccta	ccagctgctg	6480
ggccacgtct	ttccgacaga	gagctggcag	cagggtctcg	actggcgac	cacgcaagac	6540
ggccagatct	acggagacgt	catccaactg	ggcagcgccc	tcctggcccc	agccaccagg	6600
gaggcgctgg	acagatcca	cgggagcgag	ggcgcccgga	ccagctgctc	cgcgccctcc	6660
atogtcaacc	ccaacatgat	tcttgctgtc	gactctcttg	acaagttcaa	ctttacggga	6720
gccagggtgc	cgcgattoga	caccatccat	gaagagttcc	ccagggaagt	ggagtccctc	6780
gtctccttcc	cagccgactt	cttcagacca	ccggaagaaa	aagaaggccc	cctgctgagg	6840
ccggctggcc	ggaggaccac	cccgacagcc	acgcgcocgg	ggcctggcac	cgaggggag	6900
gcccgaccca	gcaaggcgag	gcgacaccc	gatgaagctg	gccagttcgc	cgctgcctgt	6960
gtcatctatt	acgcgaccc	ggggcagctc	gtcccgagc	gctacgaccc	ctgacgtctgc	7020
agcctccgtg	tgcctcaccg	gcccattatt	aatacccgga	tggtgagcac	gctggtgtac	7080
agcgagggcg	ctccgctccc	gagaccctg	gagagggcgc	tctggtgaga	gttcgcctgt	7140
ctggaggagg	aggagcgaac	caagcctgtc	tgcgtgttct	ggaaccactc	ctggccgctt	7200
ggtggagcag	gaggtgtgtc	tgcccggggc	tgctgagctc	tgccagagaa	ccggacacat	7260
gtcgcctgcc	agtgcagcca	cacagccagc	tttgcgtgtc	tcattggatat	ctccaggcgt	7320
gagaaacggg	aggtgtccgt	tctgaagatt	gtcacctatg	ccctgtgttc	ctgttcaatg	7380
gcagccctgc	tgagtgacct	cgtcctcctg	agcctggtcc	gcattgtgcg	ctccaaactg	7440
ccagcagctc	acagcaacct	cgccgtggcg	ctctctcctc	ctcagctggt	gttcctgtat	7500
gggatcaacc	agacggaaaa	cccgtttctg	tgcaagctct	ttgccatcct	ctccactcac	7560
attctacata	gcacctttgc	ctgggacccc	gtggagagcc	tgcattgtcta	ccgcatgctg	7620
accgaggtgc	gcaacatoga	caogggggccc	atggggttct	actacgtcgt	ggggctgggg	7680
atcccgccca	ttgtcacagg	actggcggtc	ggcctggacc	cccagggtta	cggggaacccc	7740
gaactctgct	ggctgtctgt	tcaagacacc	ctgatttgga	gctttcggg	ctccactcga	7800
gctgttataa	tcatcacaac	agtcactctc	gtcctatctg	caaaggtttc	ctgcocaaaga	7860
aagcaacctt	attatgggaa	aaaagggtac	gtctccctgc	ttaggacccg	atctcctcgt	7920
ctgtctgtcat	tcagcgccac	ctggctgctg	gggtcgtcgt	ctgtgaacgg	cgatgcaact	7980
agctttcaat	acctcttcgc	catcttcagc	ggcttacagg	gccccttcgt	ctctcctttc	8040
cactgcgtgc	tcaaacaggga	ggtccgggaag	caactgaagg	gcgtctcgg	cggggaagag	8100
ctgcacctgg	aggaactccgc	caccacagg	gcacccctgc	tgacgcgtc	ctccactcgt	8160
aaacccacct	ctcgtgacgg	gcttgacatg	ctgcgcacag	acttggcgga	gtccaccgcc	8220
tcgctggaca	gcattcgtcag	ggatgaagg	atccagaagc	tcggcgtgtc	ctctgggctg	8280

gtgaggggca	gccacggaga	gccagacgcg	tccctcatgc	ccaggagctg	caaggatccc	8400
cctggccagc	attccgactc	agatagcgag	ctgtcccttg	atgagcagag	cagctctctac	8460
gcctctcac	actcgtcaga	cagcaggagac	gatggggtgg	gagctgagga	aaaatgggac	8520
ccggccaggg	gcgccgtcca	cagcaccccc	aaagggggacg	ctgtggcccaa	ccacgttccg	8580
gccggttgcg	ccgaccagag	cctggctgag	agtgcagctg	aggaccccag	cggcaagccc	8640
cgctgaagg	tggagaccaa	ggtcagcgtg	gagctgcacc	gcgaggagca	gggcagctcac	8700
ggtggagagt	accccccgga	ccaggagagc	ggggggcgag	ccaggcttgc	tagcagccag	8760
ccccccagagc	agagggaagg	catcttggaaa	ataaagatca	cctacccgcc	gccgctgcag	8820
ctgacgagagc	agacgctgaa	gggcccgtctc	cgggagaagc	tggccgactg	tgagcagcag	8880
cccacatctct	cgcgcacgtc	tccctgtggc	tctggcgccc	ccgactgcgc	catcacagtc	8940
aagagccctg	ggaggggagcc	ggggcggtgac	caectcaacg	gggtggccat	gaatgtgcgc	9000
actgggagcg	cccagggccga	tggctccgac	tctgagaana	cgtgaggcaa	gcccgtcacc	9060
ccacacagcg	tgcggcatca	ccctcagacc	tggagcccca	aggggccact	gcccctgaag	9120
tggagtggcg	ccagagtgtg	gcggtcccca	tgggtggcag	ccccgactg	atcatccaga	9180
cacaaggctc	ttggttctcc	caggagctca	gggctgtgca	gacctggtga	caagtgccta	9240
aggccacagg	catgaggggag	gcgtggacca	ctgggcccag	accgctgagt	cctaagactg	9300
cagtcaaaag	cagaaactgag	aggggacccc	agactgggccc	cagaggctgc	caggagttca	9360
ggaacgcggg	gcacagacca	aagaccgcgg	tccagcccgc	cccaggccggg	catctcatgg	9420
cagtcggcac	ccgtggtctg	cagcccgggc	agtccttttg	aaaggcaccc	cttgtcttaa	9480
aatcactctg	ctatgtggga	aaggtggaga	tacttttata	taattgtatg	ggactctgag	9540
gaggtgcaac	ctgtatatat	attgcattcg	tgtgactctt	gttatcccca	gagatccatg	9600
caatgatctc	ttgtgctctt	ctctgtcaag	attgcacagt	tgtacttgaa	tttggcaatg	9660
gttgacgaaa	ctggtgcccc	agcagatcaa	aggtgggaaa	tacgtcagca	gtggggctaa	9720
aaccacagcg	ctagaagccc	tacagctgcc	tccggccagg	aagtggagat	gggtgtggcc	9780
ctcccccgcg	gccccctggg	tcccagctgt	tccgtgtgtg	tgcgtttgtc	ctctgctgcc	9840
atctgcccgc	gctgtgtgaa	tccaagacag	ggcagtgca	cactaggcag	gctgtgaggaa	9900
ccctgctgag	gtcactgttg	ggcacggttg	ccacacggct	gtcatttttc	acctgggtcat	9960
tctgtaccaa	ccaccccctc	ccctcaccgc	ctcccagggt	gcggcggaag	tgcaggtggg	10020
gatggctttg	tcctttgtct	ctgctcccgc	tgggaccttg	gaccttaag	cggtgcaggt	10080
tcctgatattg	gacacagggt	tggggccttc	caggccctta	catacctcct	gccaatctct	10140
taactctctg	agactgcgag	gatctccagg	cagggtttct	ccctctggag	tctgaccaa	10200
tacttccatt	tgcctcaaat	ggccaattgt	gcagagggac	aaagccacag	ccacactctt	10260
caacggttac	caaaactgtt	ttggaaattc	acaccaaggt	cgggcccact	cgaggcagct	10320
ggcacagcgt	ggcccagagg	gctgtggaac	gggtcccaga	actgtcagac	atgtttgatt	10380
ttagcgtttc	ctttgttctt	caaatcaggt	gcccaataa	gtgatcagca	cagctgtctc	10440
caaataggag	aaaccataaa	ataggatgaa	aatcaagtaa	aatgcaaa	tgctccacct	10500
gtttttaaact	tgacctgat	gaaaattgtga	gcactgttag	cagatgccta	tgggagagga	10560
aaagcgtatc	tgaanaatggt	ccaggacagg	aggaatgaat	gagatccag	agtcctcaca	10620
cctgaatgaa	ttatacatgt	gccttaccag	gtgagtgttc	tttcgaagat	aaaaaactct	10680
agtcctctta	aacgtttgcc	cctggcgttt	cctaagtacg	aaaagggttt	taagtctctg	10740
aacagtctcc	ttctcatgag	ttaacaggat	tctgcccctc	gaggttgat	ttttttgttc	10800
tatttttttc	cacgtactcc	acagccaaca	tcacgaggtg	taatttttaa	tctgtcaga	10860
actgttacca	aaaaacaact	gtcagtttta	ttgagatggg	aaaaatgtaa	acctattttt	10920
attacttaag	acttttatgg	agagatttaga	cactggaggt	ttttaacaga	acgtgtattt	10980
atatactgtc	aaaaacactg	aattacaatt	gagaagagtc	tacaataaat	taagattttt	11040
gaatttgtac	ttctgcggtg	ctggtttttc	tccacaaaca	cccccgcccc	tcccactgcc	11100
cagggtggcc	tgggaaggga	cggttttacg	acgtgcagct	gagctgtccg	tgctccatgc	11160
tcctcagccg	agtggaaact	gcgggaactt	tttgtccatt	ccctagttag	cctgccacag	11220
cctagatggg	cagattttgt	ctttacccaa	atttgaggac	tttttttttt	tgccattatt	11280
tcttcagttt	tcttttcttg	cactgatctt	tctctctccc	tctgtgact	ccagtgtact	11340
agacgttaga	cctcttgatg	ttttcccaact	ggtccctgag	gctctgttc	11389	

<210> 238

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014246

<400> 238
gggagagatt agacactgga ggtttttaac agaactgtga ttattattaatg ttcaaaacac 60

<210> 239
<211> 4372
<212> DNA
<213> Homo sapiens

<300>
<308> NM_014314

<400> 239
tagttattaa agttcctatg cagctccgcc tcgctccgg cctcatttcc tgggaaaaac 60
cctgttttcc ccgctgcgca cgcctcctcc ctaccggct ttaaagctag tggagcacag 120
cctgcgggga acgtagctag ctgcaagcag aggcgggcat gaccaccgag cagcgacga 180
cgctgcaagc cttccaggat tatatccgga agacccttga cctacctac atcctgagct 240
acatggcccc ctggttttag gagggaagagg tgcagtatat tcaggctgag aaaaacaaca 300
aggggccaat ggaggctgcc acacttttcc tcaagttcct gttggagctc caggaggaag 360
cgtggttccg tggctttttg gatgccttag accatgcagg ttattctgga ctttatgaa 420
ccattgaaag ttgggatttc aaaaaaattg aaaaatttga ggagtataga ttacttttaa 480
aacgtttaca accagaattt aaaaaccagaa ttatcccaac cgtatcatt ttctgactgt 540
ctgaatgttt aattaatcag gaatgtgaag aaattctaca gatttgcct actaagggga 600
tgatggcagg tgcagagaaa ttggtggaat gccttctcag atcagacaag gaaaaactggc 660
ccaaaacttt gaaactgtct ttggagaag aaaggaacaa gtcagtga cgtggattg 720
taggaaaagg tataaaagat gttgaaacag aagatcttga ggataagat gaaactctg 780
accatacagat ttcttaccaa gaagatccag aatgccagaa tcttagtgag aattcatgtc 840
caccttcaga agtgtctgat acaaaactgt acagccatt taaaccaaga aattaccaat 900
tagagtgtgc ttgtcctgtat atgaaaggaa aaaaacaaat aatattgtct cctacaggtt 960
gtggaaaaac ctttgtttca ctgcttatat gtgaacatca tcttaaaaaa ttcccacaag 1020
gacaaaaggg gaaagtgttc ttttttgcga atcagatccc agtgtatgaa cagcagaaat 1080
ctgtattctc aaaaactctt gaaagacatg ggtatagagt tacaggcatt cttggagcaa 1140
cagctgagaa tgcctcagtg gaacagattg ttgagaacaa tgacatcatt attttaactc 1200
cacagattct tgtgacaac cttaaaaagg gaacgattcc atcactatcc attcttaact 1260
tgatgatttt tgatgaatgc cacaacacta gtaaaacaac cctgtacaat atgatcatgt 1320
ttaattatct agatcagaaa cttggaggat cttcaggccc actgcccag gtcattgggc 1380
tgactgcctc ggttggtgtt ggggatgcca aaaaacaga tgaagcctg gattatattc 1440
gcaagtgtgt tgcctctctt gatgcgtcag tgatagcaac agtcaaacac aatctggagg 1500
aacctggagc agtgttttat aagccccaga agtttttcag gaaagtggaa tcacggatta 1560
cgacacaaat taaatcatc atagctcagc tgatgggga cacagagagt ctggcaaga 1620
gaactctgca agactctgaa aacttatctc aaatctcaaa tagggaattt ggaaacaga 1680
aatatgaaca atggattgtt acagtctcag aagcatgcat ggtgttccag atgcagaca 1740
aagatgagca gagcaggatt tgtaaaagccc tgtttttata cacttcacat ttgcggaaat 1800
ataatgattc cctcattatc agtgagcatg cagcaatgaa agatgctct gattacttga 1860
aagactctct cagcaatgtc cgagcagcag gattcgatga gtttgagcaa gattcttact 1920
agagatttga agaaaagctg caggaaactag aaagtgttcc cagggatccc agcaatgaga 1980
tctctaaact tgaagacctc tgcctcatct tacaagaaga gtaccactta aaccagaga 2040
caataacaat tctctttgtg aaaaaccagag cactgttggga cgtcttaaaa aattgattg 2100
aaggaaaacc taaactcagt ttctctaaaac ctgtgcattt gctggagctt ggcaaaaaca 2160
atcagaacac aggaatgacc ctccggcac agaagtgtat attggatgca ttcaagacca 2220
gtggagcaat caatatcct attgccacct cagttgctga tttgtgctga tgaaggcatt gacattgcac 2280
agtgaactct tgtcatcctt tatgagtatg tgggcaatgt catcaaaatg atccaaacca 2340
gaggcagagg aagagcaaga ggtagcaagt gcttctctct gactagtaat gctggtgtac 2400
ttgaaaaaga acaataaacc atgtacaag aaaaatgat gaatgactct attttacgc 2460
ttcagacatg ggacgaagca gtatttaggg aaaaagttct gcataacag actcatgaa 2520
aattcatcag agatagctga gaaaaaccaa aacctgcag tgaagaagg aaataaaaa 2580
tgctctcag agagtgcaaa cccttggtcat gttacacagc tgactgaaga gtgagagg 2640
aatgcocata cacgtgtcct ggagatgctt ttaaggaaat ctttgtgatt agaccacatc 2700
ccaaagccaa gcaattttca agttttgaaa aaagagcaaa gatattctgt gcccgacaga 2760
actcgacca tgaactggga atccatgtga agtacaagac atttgagatt atttgataa 2820
aaattgaaag ttttgtgtg gaggatattg caactggagt tcagacactg tactcgaagt 2880
ggaaggactt tcaattttag aagataccat ttgatccagc agaagtgtcc aatgatattc 2940

```

aggtcctcaa tcttcagcta cagggaatga gtaactttga gtggagaaga aacaaacata 3000
gtgggtataa tcatggatcg cttgtacccc tgtgaaata tattttttaa aaatatcttt 3060
agcagtttgt acatataatat atatgaaaag cacaatatgag tgaatcacag cactgagtat 3120
ttttagagcc aacagagctc atagtacttg ggaaaaataa aaagccctca tttctagctt 3180
tctttttaga gtcaactggc acaaaacaca cagtaatcac tctgtacaca ctgggataga 3240
tgaatgaatg gaatgtgtgg aatttttatc tccctttgtc tcttaaacct acgtgaaact 3300
ggcttttgcc cttaaacaat tactgaaatt gttcttttga aggttaccag tgactctggt 3360
tgccaaatcc actggggcact tcttaacott ctatttgacc tctggcatt tggccctggt 3420
gagcaotctt cttgaagctc tccctgggct tctctctctt ctagtcttat tctagtcttt 3480
ttttattgag tccctctctt tgctgatccc ttccaagggt tcaatatata tacaatgata 3540
tactgtacat atgtatatgt aactaatata catacataca ggtatgtata tgtaatgggt 3600
atatgtactc atgttctctg ttagcaacog tgtggtatgg ctacacagag aacatgagaa 3660
cataaagcca tttttatgct tactactaaa agctgtccc tgtagagttg ctgtatgtag 3720
caatgtgtat ccaactgtaca gtggtcagct ttagtagag agcataaaaa tgataaaaaa 3780
ctctctgaaa acttatgtta ctatacatct tgccctatta atagtctctc ttaactgtgt 3840
ccattgttct ctttgacctt tttcctataa tgatgttgat gttcaacacc tggactgaat 3900
gtctgttctc agatcccttg gatgttacag atgaggcagt ctgactgtcc tttctacttg 3960
aaagattaga atatgtatcc aaatggcatt cagctgtcac ttagcaagggt ttgacgagtg 4020
ttcaaaaagc ttagttttgg gtttctctga cgtgaaaaca agtatctag ttccctggag 4080
atcaaacagg tgaagtgtaa cagctgcctc cctctctcat caatctgggt agcagtggtg 4140
caggcgggga gccagagaaa cttgccagtt atataacttc tctttggctt tctctcatct 4200
gtaaaaaag gataatactg aactgttaagg gttagtggag agtttttaat taaaagaatg 4260
gtgaaaaagt acatgacaca tagtttgctt gataataagt actagtagta gtttcttacc 4320
taagacccaa tacaatgga ttatttaaac caaaaaaa aa 4372

```

<210> 240

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014314

<400> 240

agtcagaca ctgtactcga agtggaaaga ctttctattt gagaagatac catttgatcc 60

<210> 241

<211> 1647

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014321

<400> 241

```

gcgcgcgggt ttcgttgacc cgcggcggtc acgggaattg ttcgcttttg tgcggcgccc 60
atgggggtcgg agctgatcgg gcgcctagcc cgcgcgcttg gcctcgccga gcccgagcatg 120
ctgaggaaagc cagaggagta cttgcgcctg tcccggggtga tgtgtgtcgg cctcccgca 180
cgccacccagg agaccagcag tgcagtcatg tgcctggacc ttgcagcttc ctggatgaag 240
tgcctccttgg acaggcctta ttttaattaa ggcctgaatt tgaacaagga gacatatcag 300
agctgtctta aatcttttga gtgtttactg ggcctgaatt caaatattgg aataagagac 360
ctagtgttac agtttagctg tatagaagca gtgaacatgg ctccaagat actaaaaagc 420
tatgagtcga ctcttcccca gacacagcaa gtggatcttg acttatccag gccacttttc 480
actcttctgt agctgcttcc agcatgcaag attctaaagg tgaagtggg taataacaaa 540
atggtgacca catccgtgtg aaaaaaacgt atatttgatc gactgtgtaa acaactagag 600
aaagttgacg agcaggtcga cagagaacct ctactccacc acggaagaga 660
aagaagatag tggttgaagc ccagcagaag gaaatggaga aggtagagga gatgccacat 720
aaaccaagca aagatgaaga tctgacacag gattatgaag aatggaaaag aaaaattttg 780
gaaaatgctg ccagtgtcca aaaggctaca gcagagtgat ttcagcttcc aaactggtat 840
acattccaaa ctgtagtagc attgccatct ccaggaaagc ttgacgctt tgggattttg 900

```



```

tttaaacctt tataataagg atcctaagac tgttgcttt aaatagcaaa gcagcctacc 960
tggaggctaa gtctggggac tgggctggcc cctgggtgga gcattagacc agccacagt 1020
cctgattggt atagccttat gtgctttcct acaaaatgga attggaggcc ggccgcagt 1080
gctcaccgct gtaactccag cactttggga ggccaagggt ggtggatcac ctgaggctag 1140
gagctcgaga ccagcctggc caacatgggt aaccccatc tctactaaaa atacaaaaat 1200
tagccagggt tgaatgggca tgccctgtaat ccagctcctc cagtaggctg agacaggagg 1260
atcacttgaa cgtggggagg agagggtgca gtgagccgag attgcaccac cgcactccag 1320
cctgggtgac agagcggagac ttatctcata aataaataga tagatactcc agcctgggtg 1380
acagagcgag acttatagat agatagatag atagatggat agatagatag atagatagat 1440
agatagataa accgaatttg agccattttg ctttaagtga atggcagctcc ctgtgtctat 1500
tcagaataata aaattcagtc tgaatggcat cttacagatt ttacttcaat ttttgtgtac 1560

```

```

ggtatttttt atttgactaa atcaatatat tgtacagcct aagttaataa atgttattta 1620
tatatgcaaa aaaaaaaaaa aaaaaaa 1647

```

```

<210> 242
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_014321

```

```

<400> 242
tgctttaagt gaatggcagt cccttgtctt attcagaata taaaattcag tctgaatggc 60

```

```

<210> 243
<211> 1455
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_014364

```

```

<400> 243
ggcgggtccg acgcacctcg gtaacatcac agcagggtcca ggccaatgat aacctataaa 60
gaggccatgt cgaagcgcca catcgctctc accaatgtca ccgttgtcca gttgctgcga 120
cagcgtgccc cggtgaccag agcacgcgcc ccacctgagc ctaaggctga agtagagccc 180
cagcccaaac cagagccccc accagtcagg gaggaaataa agccaccacc gccaccactg 240
cctcctcacc ccgctactcc tctcctaaag atggtgtctg tggcccgagg gctgactgtg 300
ggcataaatg gatttggacg catcggtcgc ctggtctcgc ggcctctcat ggagaaaggt 360
gttaaggtgg tggctgtgaa tggatccattc attgaccggg aatacatggt gtacatgttt 420
aagtatgact ccacccacgg cagatacaag ggaagtgtgg aattcaggaa tggacaactg 480
tcgtgggaca accatgagat ctctgtctac cagtgcgaag agcccaaaaca gatccctgg 540
agggctgtcg ggaagcccta cgtgtgtggg tccacaggcg tgtacctctc catcacggca 600
gcttcggacc acatctctcg aggtgctcaa cgtgtgtgca tctccgccc ctcacgggat 660
gcaccaatgt tctgcatggg tgtcaatgaa aatgactata accctggctc catgaacatt 720
gtgagcaacg cgtcctgcac caccacatgt ttggctcccc tcccaaaagt catccacgag 780
cgatttggga tctgtggaag gttgatgacc acagtccatt cctacacggc accccagaag 840
acagtggaag gcccatcaag gaaggcctgg cgagatgggc ggggtgccca ccagaaacat 900
atccagacct ccactggggc tgcgaaagct gtgacaaaag tcatccaga tctcaagggt 960
aagctgacag ggaatggggt ccgggtacca accccggatg tgtctgtcgt ggacctgacc 1020
tgccgcctcg ccacagcctgc cccactactca gccatcaagg aggcgtgaaa agcagcagcc 1080
atcggggccc ttgctggcat ccttgcttac accgaggatg aggtcgtctc tacggacttc 1140
cctgtgtgata cccactcgtc catcttcgat gctaaggccg gcattgcgct caatgacaat 1200
tctgtgaagc tcaatttcatt gtaacgcaac gaatatggct acagtccagg ggtggtcgac 1260
ctctccctcg acatgttcag ccgagacaag tgaaacggga aggtctcttc ttctcttccc 1320
agggcgccgg gccggaacat gtgcctcccg ttccagcatc tggctgcccg ggggaggaag 1380
gacaccgggg gcggcgccc caccggcatg ggtccatggt gaaataaaaa acagtgcctg 1440
aaaaaaaaa aaaaa 1455

```

<210> 244
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014364

<400> 244
 cgctcaatga caatttcgtg aagctcattt catggtacga caacgaatat ggctacagtc 60

<210> 245
 <211> 935
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014462

<400> 245
 gaagtgggtg agggtaatat ggaggagctt cgggcaggcc cggcgggctg aaagccgggg 60
 cagaagtgcg ggtctcggtc gggattccgg gcttggtccc accgaggcgg cgactgcggt 120
 aggggggaag aggttttggg cgcgctggcc tccgcgcgct gtgcattgca gcattatttc 180
 agttcaaaat gaactatatg cctggcacgc ccagcctcat cgaggacatt gacaaaaagc 240
 acttggttct gcttcgagat ggaaggacac ttataggctt ttaagaagc attgatcaat 300
 ttgcgaactt agtgcacat cagactgtgg agcgatttca tgtgggcaaa aaatacgggtg 360
 atattcctcg agggattttt gtggtcagag gagaaaatgt ggtcctacta ggagaaatag 420
 acttggaaaa ggagagtgac acacccctcc agcaagtatc cattgaagaa attctagaag 480
 aacaaagggt ggaacagcag accaagctgg aagcagagaa gttgaaagtg caggccctga 540
 aggcgcgagg tctttccatt cctcgagcag atactcttga tgagtactaa tcttttgccc 600
 agagggtggt ggctcttgaa gagtaggggc tgtcactgag tgaagtgcac atcctggcca 660
 cctcagcagc ttgatcacag actgtagagt ttgaaaagt cacttttatt ttaattatt 720
 ttacatatgc aacatgaaga aatcgtgtag gtgggttttt ttttaataa caaaatcaact 780
 gtttaaagaa acagtggcat agactccttc acacatacct gtggcaccag caactacttc 840
 tttatatgtt tcttcataac ccaaattaga gtttacaggg acagctctca tttacttgta 900
 aataaaatat gaatctcaaa aaaaaaaaaa aaaaa 935

<210> 246
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014462

<400> 246
 ttaataacaa aatcactgtt taaagaaaca gtggcataga ctcccttcaca catcactgtg 60

<210> 247
 <211> 890
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014501

<400> 247
 ggcggacgca agaacgcagg aagggggcgg gggggaccgg ccccgggcgg gccgcagcca 60
 tgaactccaa cgtggagAAC ctaccccccg acatcatcgg cctgggtgtac aaggaggtga 120
 cgacactgac cgagacacca ccgatggcga tcaaggtctt tcccaacgag gaggacctca 180
 ccgacctcca ggtcaccatc gagggccctg aggggacccc atatgctgga ggtctgttcc 240

```

gcacgaaact cctgctgggg aaggacttcc ctgcctcccc acccaagggc tacttctcga 300
ccaagatctt ccaccocgaac gtggcgccca atggcgagat ctgcgtcaac gtgctcaaga 360
gggaactggc ggctgagctg ggcatccgac acgtactgct gaccatcaag tgcctctga 420
tccaccctaa ccccgagtct gcaactcaac agggaggcggy ccgcctgctc ttggagaact 480
acgaggagta tggcggtctgg gccgctctgc tcacagagat ccacggggggc gccggcggggc 540
ccaggcgagc ggccgaagcc ggtcgggccc tggccagtgg cactgaagct tectccacccg 600
acctcggggc cccagggggc ccgggagggg ctgagggtcc catggccaag aagcatgctg 660
gcgagcgcgga taagaagctg gcggccaaga aagaagcgga caagaagcgg gcgctcgggg 720
cgctcgggcg gctgtagtgg gctctcttcc tccttccacc gtgaccccaa cctctcctgt 780
ccctccctcc caactctgct tctaagttat ttaaatatg gctgggggtac 840
agggggcact gggacctgga tttgtttttc taaataaagt tggaaaagca 890

```

<210> 248

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014501

<400> 248

```

acacgtactg ctgacctca agtgctctgt gatccacct aaccocgagt ctgcactcaa 60

```

<210> 249

<211> 1182

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016095

<400> 249

```

gcgccggggc gcgtctcttc cggggacgct gaggggcccg aggagacgt gaggctctgg 60
cctgcagctc ggcggcccat ggacgctgcc gaggtcgaat tctcgcgca gaaggagctg 120
gttaccatta tccccaactt cagtctggac aagatctacc tcactcgggg ggacctgggg 180
ccttttaacc ctggtttacc cgtggaaagt cccctgtggc tggcgattaa cctgaaacaa 240
agacagaaat gtgcctgct cctccagag tggatggatg tagaaaaatt ggagaagatg 300
agggatcatg taagaaaagg aaaaactttt accccaatgc ccagccctta ctacatggaa 360
cttacgaagc tctgtttaa tcatgtctca gacaacatcc cgaaggcaga cgaatccgg 420
acctctgtca aggatgtg ggacactcgt atagccaaac tccgagtgtc tgcctgacgc 480
tttgtgagac agcaggaggc acatgccaa gtgataact tgaccttgat ggagatcaac 540
accagcgagg ctttctcac acaagcgctc aaccacatg acaactccg ccggaacctc 600
cagcctctgg agagtactca gtctcaggac ttctagaaa aggcctgggt caggcgctgt 660
gctgggggat gtgagcgtc aggatgtgat gaggctactc tggttctgga gctctagaaa 720
cactctgatg gcatgaaaaa tgtgtgatgg tgcaggaaat ggattcagga tgttgttga 780
gaacacaagt tgtgattagt ccttaaaact tagctccctg ggaactcttt caattccaca 840
ttcgtttcta gaaccagcc ctttttcccc ccaactttga gaaataaaaa agccttaggt 900
aaataagtca ttctccctag cagagccaat tgggtctcct gcatgggaag cgtcacactt 960
gggcaggtgt tcaagtgaatg taggtgttag atacagcagg agtgggccatg tgggtccagg 1020
ctttttaccc cttcttgatc ctgatttctt gggctgaatt tagactctct cacagaggtg 1080

```

```

gtccacagag aaggatggca gatggtgcag ccaacaatgc tgaccgggtc ttatcctcta 1140
agccctgatc cacaataaaa atggacccaa ctcaaaaaaa aa 1182

```

<210> 250

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016095

<400> 250
atggattcag gatgttggtg gagaacaag tttgtgatta gtccttaaaa cttagctccc 60

<210> 251
<211> 704
<212> DNA
<213> Homo sapiens

<300>
<308> NM_016185

<400> 251
tgcagcgggt gtcgcgtggt ggggtgtggag ttccccagcg cccctcgggt ccgacccttt 60
gagcgttctg ctcgcgcgcc agcctacctc gctcctcggc gccatgacca caaccaccac 120
cttcaaggga ctcgacccca acagcaggaa tagctccgga gttttcgggc ctcaggttgg 180
tggatccaat ttttcattag gttttgatga accaacagaa caacctgtga ggaagaacaa 240
aatggcctct aatatctttg ggacacctga agaaaaatcaa gcttcttggg ccaagtccgc 300
aggtgccaag tctagtgggt gcagggaaga cttggagtca tctggactgc agagaaggaa 360
ctcctctgaa gcaagctccg gagacttctt agatctgaag ggagaagggt atattcatga 420
aaatgtggac acagacttgc caggcagcct ggggcagagt gaagagaagc ccgtgctctg 480
tgcgctctgt cccagccggc tggcccgggc ccagtgcca tccagaagaa atccccctgg 540
cggcaagtcc agcctcgtct tgggttagct ctgactgtcc tgaacgctgt cgttctgtct 600
gtttcttcca tgcttgagaa ctgcacaaact tgagcctgac tgtacatctt cttggatttg 660
tttctataaa aagaagcact ttatgtataaa aaaaaaaaaa aaaa 704

<210> 252
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_016185

<400> 252
tgaaccaaca gaacaacctg tgaggaagaa caaaatggcc tctaatactt ttgggacacc 60

<210> 253
<211> 2268
<212> DNA
<213> Homo sapiens

<220>
<221> Modified_base
<222> 1 ... 2268
<223> n = a, c, g, or t

<300>

<308> NM_016359

<400> 253
gggatttgaa ccncgcgtgc gaagtttggt gatccatctt ccgagtatcg ccgggatttc 60
gaatcgcgat gatcatcccc tctctagagg agctggactc cctcaagtac agtgacctgc 120
agaacttagc caagagctctg ggtctccggg ccaacctgag ggcaaccaag ttgttaaaag 180
ccttgaaagg ctacattaaa catgaggcaa gaaaaggaaa tgagaatcag gatgaaagtc 240
aaaacttctgc atctcttctg gatgagactg agatacagat cagcaaccag gaagaagctg 300
agagcagacc acttgcccat gtcacaaaaa caaggagaag gtgcaagact gtcctgtgtg 360
acctgactc acagcagaat cattcagaga taataataag taatcccatc gaattccaga 420
atcatgaaaa gcaggaaagc caggatctca gagctactgc aaaagttcct tctccaccag 480
acgagcacca agaagctgag aatgctgttt cctcaggtaa cagagattca aaggtacctt 540

cagaagaaa	gaaatctctc	tacacagatg	agtcacccaa	acctggaaaa	aataaaagaa	600
ctgcacacac	tactccaaac	tttaagaagc	ttcatgaagc	tcatttttaag	gaaatggag	660
ccattgataca	atatattgag	agaaaaagaa	acattttgaa	gaacacaatt	ccatgaatga	720
actgaagcag	cagcccatca	ataaggagg	ggtcaggact	ccagtacctc	caagaggag	780
actctctgtg	gcttctactc	ccatcagcca	acgacgctcg	caaggccggt	cttgtggccc	840
tgcaagtccag	agtaccttgg	gtctgaagg	gtcactcaag	cgctctgcta	tctctgcagc	900
taaaacgggt	gtcaggtttt	cagctgctac	taagataat	gagcataaag	gttcaactgac	960
caagactcca	ggcagaagt	ctgcacatgt	gaccgtgtct	ggggggcacc	caaaaggcga	1020
ggctgtgctt	gggacacaca	aattaagac	catcacgggg	aattctgctg	ctgttattac	1080
ccattccaag	ttgacaactg	aggcaacgca	gactccagtc	tcacaataaga	accagtggtt	1140
tgatcttaaa	gcaagtgtgt	ctcgtccctc	caactatgaa	ccacacaag	gaaagctaaa	1200
accatggggg	caatctaaag	aaaataatta	tctaatacaa	catgtcaaca	gaattaaact	1260
ctacagaaga	acttacaac	aacccatct	ccagacaag	gaagagcaac	ggaagaaacg	1320
cgagcaagaa	cgaaggaga	agaaagcaaa	ggttttggga	atgcgaagg	gcoctatttt	1380
ggctgaagat	taataatttt	ttaatatctt	gtaaatattc	ctgtattctc	aacttttttc	1440
tttttctaaa	tttttttttt	tttgcgtgca	tcccaacttt	agtcacgaga	tctttttctg	1500
ctaaactgtt	atagtctgtg	tagtgtccat	gggtctctca	tgtgctatga	tctctgaaaa	1560
gacgttaatca	ccctaaagct	caaattcttt	gggatgggtt	ttacttaagt	ccattaaaca	1620
ttcaggtttc	taacgagacc	catcctaaaa	ttctgtttct	agatttttaa	tgtcaagtgc	1680
ccaagtctcc	ctgtctggtt	ctaataat	cagaactgca	gtctctgtct	agccaatagc	1740
atttacctga	tggcagctag	ttatgcaagc	ttcaggagaa	tttgaacaat	aacaagaata	1800
gggttaagctg	ggatagaag	gccacctctt	caactctcat	agaataatag	aacctttatg	1860
aaacggggcc	ataatgtttg	gttatgacat	caatatatta	cctagtgtaa	atgttttagg	1920
cttatgtacc	ttcgttcaaa	tatcctcatg	taattgccat	ctgtcaactca	ctatattcac	1980
aaaaataaaa	ctctacaact	cattctaaac	ttgtctactt	aaaagctcac	tattgcctatc	2040
gaaatgcgag	gattaatgct	ttaatgcttt	tagagacag	gtctcaactgt	gttgcccagg	2100
ctggtctcaa	actccaccaa	atgtactctt	tattctattt	atggaaaaga	ctaggctttg	2160
cttagtatca	tgtccatggt	tccttcacct	cagtggaagt	tctgagtttt	atactgctca	2220
agatcgtcat	aaataaaatt	ttttctcatt	gtcaaaaaaa	aaaaaaaa	2268	

<210> 254

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016359

<400> 254

acattgctta cttaaaaagct acatagccct atcgaaatgc gaggattaat gctttaatgc 60

<210> 255

<211> 1590

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016816

<400> 255

gaggcagttc	tgttgccact	ctctctcctg	tcaatgatgg	atctcagaaa	taccccagcc	60
aaatctctgg	acaagttcat	tgaagactat	ctcttgcag	acacgtgttt	cgcgatgcaa	120
atcgacacatg	caattgacat	catctgtggg	ttcctgaag	aaaggtgctt	cagaggtagc	180
tcctaccctg	tgtgtgtgtc	caaggtggta	aagggtggct	cctcaggcaa	gggcaccacc	240
ctcagagccc	gatctgacgc	tgacctgggt	gtcttctca	gtcctctcac	caactttcag	300
gatcagttaa	atcgccgggg	agagttcatc	caggaaatta	ggagacagct	ggaagcctgt	360
caaagagaga	gagcactttc	cgtgaagttt	gaggtccagg	ctccacgtct	ggcacaaccc	420
ctgtgcctca	gcttcgtact	gagttcgctc	cagctcgggg	aggggggtga	gttcgatgtg	480
cgtgcctgct	ttgatgccc	gggtcagttg	actggcagct	ataaacctaa	cccccaaatc	540
tatgtcaagc	tcacgagga	gtgcaccgac	ctgcagaaag	agggcgagtt	ctccacctgc	600
ttcacagaac	tacagagaga	cttctcgaag	cagcgcccca	ccaggtctca	gagcctcatc	660

cgccatgta	agcactggta	ccaaaattgt	aagaagaagc	ttgggaagct	gccacctgag	720
tatgcctgg	agctcctgac	ggctcatgct	tgggagcgag	ggagcatgaa	aacacatttc	780
aacacagccc	aaggatttcg	gacggtcttg	gaattagtca	taaactacca	gcaactctgc	840
atctactgga	caaagtatta	tgacttttaa	aacccatta	ttgaaaagta	cctgagaagg	900
cagctcacga	aaccagggcc	tgtgatcctg	gacccggcgg	accctacagg	aaacttggtt	960
gggtggagacc	caaaggggtg	gaggcagctg	gcacaagagg	ctgaggcctg	gctgaattac	1020
ccatgcttta	agaattggga	tgggtcccca	gtgagctcct	ggattctgct	ggctgaaagc	1080
aacagtagac	acgatgagac	cgacgatccc	aggaagctatc	agaaatatgg	ttacatggga	1140
acacatgagt	accctcattt	ctctcataga	ccagcagcgc	tccaggcagc	atccacccca	1200
caggcagaag	aggactggag	ctgcaccatc	ctctgaatgc	cagtgcactc	tgggggaaag	1260
ggctccagtg	ttatctggac	cagttccttc	atcttcagtg	gggactcttg	atccagagaa	1320
gacaaagctc	ctcagtgtag	tgtgtataaa	tccaagacag	aaccacagtc	tcctgactcc	1380
tggccttcta	tgcctctcat	ccatcatatg	ataacattct	ccacagcctc	acttctattcc	1440
acctattctc	tgaaaaattt	ccctgagaga	gaacagagag	atttagataa	ggaattgaaa	1500
ttccagcctt	gactttcttc	tgtgcacctg	atgggagggg	aatgtcta	gtattatcaa	1560
taacaataaa	aataaagcaa	ataccaaaaa	1590			

<210> 256

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016816

<400> 256

cgatcccaagg acgtatcaga aatatgggta cattggaaca catgagtacc ctcatcttctc 60

<210> 257

<211> 2905

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016817

<400> 257

cgagcggcag	ctgagagcaa	tgggaaatgg	ggagtcocag	ctgtctctcg	tgcctgctca	60
gaagctgggt	tgttttatcc	aggaatacct	gaagccctac	gaagaatgct	agacactgat	120
cgacgagatg	gtgaacacca	tctgtgacgt	ctgcaggaac	cccgacaagt	tccccctggg	180
cgagggagtg	gccataggtg	gctcctatgg	acggaaaaaa	gtcttaagag	gcaactccga	240
tggtagctct	gtccttttct	tcagtgactt	aaaacaattc	caggatcaga	agagaagcca	300
actgtgacat	ctcgataaaa	ctggggataa	gctgaagtgc	tgtctgttca	cgaaagtgggt	360
gaaaaacaat	ttcgagatcc	agaagtcctc	tgtatgggtc	accatccagg	tgttcacaaa	420
aaatcagaga	atctctttcg	aggtgctggc	cgctctcaac	gctctgagct	taaatgataa	480
tccagcccc	tggatctatc	gagagctcaa	aagatccttg	gataagacaa	atgcagatcc	540
tgggtgagtt	gcagtcctgt	tcaactgaact	ccagcagaag	ttttttgaca	accgctcctg	600
aaaactaaa	gatttgatcc	tcttgataaa	gcactggcat	caacagtgc	agaaaaaaat	660
caaggattta	ccctcgctgt	ctccgtatgc	cctggagctg	cttacggtgt	atgcctggga	720
acagggggtg	agaaaaagaca	actttgacat	tgtcgaaggc	gtcagaacgg	ttctggagct	780
gatcaaatgc	caggagaagc	tgtgtatcta	tgggatggct	aactacaact	tgaagatga	840
gacatcagca	aaactcctgc	tgcaccagct	ccaatcagcg	aggccagtaa	tcttgatccc	900
agttgaccca	accaataatg	tgaagtggaga	taaaataatg	tggcaatggc	tgaaaaaaga	960
agctcaaaac	tggttgactt	ctcccaacct	ggataatgag	ttacctgcac	catcttgaaa	1020
tgtcctgcct	gcaccactct	tcaacgacct	agggcaacct	ctggataagt	tcatacagga	1080
gtttctccag	cccaacaact	gcttctctaga	gcagattgac	agtgcgtgta	acatcatccg	1140

taccattcctt

aaagaaaact gcttccgaca atcaacagcc aagatccaga ttgtccgggg 1200

aggatacaac gccaaaggca cagctctgaa gactggctct gatgcgcgat tgcgtctgtt 1260

ccataactca cttaaaaagt acacctccca aaaaaacgag cggcacaaaa tctgtaaggga 1320

aatccatgaa cagctgaaag ccttttgag ggagaaggag gaggagcttg aagtcagctt 1380

```

tgagcctccc aagtggaaag ctcccagggt gctgagcttc tctctgaaat ccaaagtctt 1440
caacgaaagt gtcagctttg atgtgcttcc tgcctttaat gcactgggtc agctgagttc 1500
tggtctccaca cccagccccc aggtttatgc agggctcatt gatctgtata aatctcggga 1560
cctcccggga ggagagtttt ctacctgttt cacagtcctg cagcgaaact tcattcgctc 1620
ccggccacc aaactaaagg atttaattcg cctgggtgaag cactggtaca aagagtgtga 1680
aaggaaactt aagccaaaagg ggtctttgcc cccaaagtat gccttggagc tgcctcccat 1740
ctatgcctgg gagcagggga gtggagtgcg ggattttgac actgcagaag gtttccggag 1800
agtctctggag ctggtcacac aatatcagca gctcggcatc ttctggaaag tcaattacaa 1860
ctttgaagat gagaccgtga ggaagttttct actgagccag ttgcagaaaa ccaggcctgt 1920
gatcttggac ccaggcggaac ccacaggtga cgtgggtgga ggggacgctt ggtgttggca 1980
ttcttctggac aaagaagcaa aggttaggtt atcctctccc tgcctcaagg atgggactgg 2040
aaacccaata ccaccttggg aagtgcgcac aatgcagaca ccaggaaagt ttggagctag 2100
gatccatcct attgtcaatg agatgttctc atccagaagc catagaaatcc tgaataataa 2160
ttctaaaaag aactctctga gatcatctgg caatcgcttt taaagactcg gctcacgctg 2220
agaaagagtc actcacatcc attcttccct tgaagtctcc tattcctcct tcccttgect 2280
tcttggactct ctgtgaaatca atcaagactg caaacctctt cataaagctg ccttgctgaa 2340
ctctctctctg caggagccct gottaaaaata gttgatgtca tcaactttatg tgaactttat 2400
ttctgtcaac ttgtattttt ttttcttgta tttttccaat tagctctccc ttttctctc 2460
cagtcataaaa aaggaaatct ctgtgtcttc aaagcaaacg tctttacttt ccccttgggt 2520
ctcataaactc ttgtgtcttg ctctcgtgtc ttccaaactca tccactctct gtcctttctc 2580
ttgtatatac aaacctcttc tgcccctgct gacacagaca tccctctatg cagcagccag 2640
gcacaccttt tcattagaac tcaagctctc ccaaggctcc agattataac tgttgtcata 2700
tttatatgag gctgtgtgtc ttctctctg agcctgcctt tatccccca cccaggagta 2760
tcctcttgcc aaagcaaaag actttttctc tggctttagc cttaaaagata cttgaaggtc 2820
taggtgcttt aacctcacat accctcaact aaacttttat cactgttgca tatacaggtt 2880
gtgatacaat aaagaatgta tctgg 2905

```

<210> 258

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016817

<400> 258

aagggtctagg tgctttaacc tcacataccc tcacttaaac tttttacact gttgcatata 60

<210> 259

<211> 2054

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017414

<400> 259

```

gggaagctcg ggcgcgcaag gtttccccgc acgctggcgc ccagctcccg gcgcggaggc 60
cgctgtaagt ttgcgtttcc attcagtgga aaacgaaaagc tggcgggggt gccacgagcg 120
cggggccaga ccaaggcggg cccggagcgg aacttcggtc ccagctcgtg ccccggtcca 180
gtccgcagct ggaactcagc agcgaggctt ggacgcttgc atggcgcttg agagattcca 240
tcgtgcctgg ctacataaag cgcttctcgt aagtgaagtc gtgctgtcct gaacgcgggc 300
caggcagctg cggcctgggg gttttggagt gatcacgaat gagcaaggcg tttgggtccc 360
tgaggcaaat ctgtcagtc atcctgggct agtctcgcga gctccccgca gatcttgaag 420
aaaagaagga agaagacagc aacatgaaga gagagcagcc cagagagcgt cccagggcct 480
gggactaccc tcatggcctg gttggtttac atcaacttgg acagacctcg tgccttaact 540
cctgtattca ggtgtctgta atgaatgtgg acttcaaccg gatattgaag aggatcaagg 600
tgcccagggg agctgacgag cacagggaga gcgtcccttt ccagatgctt ctgctgctgg 660
agaagatgca ggacagccgg cagaagaagc tgcggccctt ggagctggcc tactgctgct 720
agaagtgcac cgtgccttg tttgtccaac atgatgtgcg ccaactgtac ctcaaaactct 780
ggaacctgat taaggaccag atcactgatg tgcacttggt ggagagactg caggccctgt 840

```

```

atacagcatccg ggtgaaggac tcccttgattt gcgttgactg tgccatggag agtagcagaa 900
acagcagcatc gctcacccctc ccactttctc tttttgatgt ggactcaaaag cccctgaaga 960
cactggaggga cgccctgcac tgctttctcc agcccaaggga gttatcaagt aaaaagcaag 1020
gctctctgtga gaactgtggg aagaagaccg gtgggaacaa ggtcttgaag ctgaccatt 1080
tgccccagac cctgacaatc cacctcatgc gattctccat caggaattca cagacagaaa 1140
agattctgcca ctccctgtac ttcccccaga gcttgagatt caggccagatc ctccaatga 1200
agcgagagtc ttgtgatgct gaggagcagt ctggagggca gtagagcgtt ttgtctgtga 1260
ttgcgcagct gggaaatggca gactccggtc attactgtgt ctacatccgg aatgctgtgg 1320
atggaaaatg gtctctgctc aatgactcca atatttgctt ggtgtccctg gaagacatcc 1380
agtgatcacta cggaatactc aactaccact ggcaggaaac tgcatatct ctggtttaca 1440
tgaagatgga gtgctaagtg aaatgcccaa aaccttcaga gattgacacg ctgtcatttt 1500
ccatttccgt tctctgactc acggagtctt tgcagtga tgcaatgag gagaagcatt 1560
gttttcaaac tatataactg agccttattt ataattaggg atattatcaa aatagttaac 1620
catgaggccc ctcaagtgct gatcagtcag aatggatgct ttaccagca gacccggcca 1680
tgtgtgctct cggtcctggg tgctctgctc tgtcgaagc attagccctt tagttatgag 1740
cctgtgggaa cttcacagggt tcccagtggg gagagcagtg gcagtgggg gcactctgggg 1800
gccaaaggtc agtggcaggg ggtatttcag tattatacaa ctgtctgtgac cagactgtga 1860
tactgcctga atatcagtcg tgtttgtaat ttttcacttt gagaaccaac attaattcca 1920
tatgaatcaa gtgttttcta actgctattc atttattcag caaatattta ttgatcatct 1980
cttctccata agatagtgtg ataaacacag tcatgaataa agttattttc cacaaaaaaa 2040
aaaaaaaaa aaaa 2054

```

<210> 260

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017414

<400> 260

tgagcatctc ttctccataa gatagtgtga taaacacggt catgaataaa gttattttcc 60

<210> 261

<211> 3638

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017523

<400> 261

```

ggtagatgct gctgtgacag cagcaaaagaa tgacggccaa gggcgacagc aggggctggc 60
cctgtgttaa aggggcttct tgggagggtc cagcctcagg aatcaagggg aactcctgag 120
cgagaaatc tgaagatctc ctccctccct gaagctgtgg gctggggccat cggaaaactt 180
tcagttttgt ttccctgcct gcaagaaaag aaactcaacc gaaagcctgc agagagcaga 240
acatggaagg agacttctcg gtgtgcagga actgtaaaag acatgatgtc tctgccact 300
tcaccctcca tgaggcttaac tgccctcggt tccctggtct gtgtccggag tgtgaggagc 360
ctgtccccaa ggaaccatcg gaggagcagt gcaagcttga gcaaccagca gttgggtgta 420
cgatgtgtca cgagagcatg cagaagtctc cgtggaagt tcataaggcc aatgagtgcc 480
agtagcgccc tglttgltg aagttctgca aactggaact cagctcagc agagctcagc 540
tccacgagtc ctactgtggc agcgggacag agctctgcca aggctgtggc cagttcatca 600
tgacccagct gctcgcccaag cacagagatg tctgtgcag tgaaacagcc cagctcggga 660
aaggggaaag aatttcagct cctgaaaagg aaatctactg tcattattgc aaccaaata 720
ttccagaaaa taagtatttc caccatatgg gtaaatgttg tccagactca gattttaaga 780
aacactttcc tgttggaatc ccagaaatc ttcttctcat tcttccaagt caagctcgtg 840
aaaatcaaac ttccacgatg gagaaaagtg ttctgtccaa gacaagaaat ataaacagat 900
ttctcttcca ttctgaaagt tcatcaaaag aagcaaaaac aacgaaaatc aaaaccttgg 960
atccactttt gatgtcagag cccaagccca ggaccagctc cctagagga gataaagcag 1020
cctatgacat tctgaggaga tgttctcagt gtggcatcct gcttccccct ccatccttaa 1080
atcaacatca ggagaaatgc cgttggttag ctctcatcaa aggaaaacaa gtgagaatt 1140

```



```

tcagctagat ttgaaaaagg aaaggtacta caaattcaaa agatttcaact ttaaacactg 1200
gcattctcgc ctacttgcgt ttggtggtcct gtgaaaagggt atgggttttta ttggttgggc 1260
tttaaaaggaa aaggtttggc agaactaaaa acaaaactca cgtatcatct caatagatcac 1320
agaaaaggct ttgtataaaa ttcaacttga ctctcatgta aaaacctctca caaacaccgg 1380
cgtcgaaggc acatcacaaa aaataataag agccatctat gacaaaacca cagcoaacat 1440
catactgaat gagcaaaagc tggagcatta ctcttgagaa gtgaaacaag gcacttcagt 1500
octattcaac atagtactgg aagtcctcgc cacagcaatc agggcaagaga aagaaataaaa 1560
aggcaaccaaa aaagaaaggc agtcgaagta tctctgtttg cagacgatat gattctatat 1620
cttgaaaacc ccattgactct ggcccaaaag ctctcatgac tgataaacaa ctccagctaa 1680
cttagaggag ccaaaatcaa tatcaaaaat atggttagcat ttttatcac ccaacgacat 1740
caagctgaga gccaaatcaa gaatgcaatc ctattcacaa ttggccacaaa aagaataaaa 1800
catttaggaa tacagctaac caggagatg aaagatctct acaacaaaaa ttgcaaaaaa 1860
ctgctgaag aaatcagaga tgacacaaat ggaaaaaat tccatcatta tggataggaa 1920
gaatcaatat tgttaaaat gccatctac ccaaaagcaat ttatagattc aatgctattc 1980
ctatcaact accataaaca ttcttcacag aatcagaaaa aaaaagcatt aaaatttatt 2040
tgaaaaccaaa aaagagccca aaaaagccaaa gcaatcctaa gcaaaaaaaa caaagctgga 2100
ggcatcgcat taccacactt caaactatcac tacagggcta cagtaaccaa aactgcattga 2160
tactggtaca aaagcatggt gctggttaca aagcagacac atagatcaat ggaacagaat 2220
agagggccca gaataaaagc tacacaccta caaccatcta atctttgaca aagttgacaa 2280
aaatcgcgaa tggggaaaga atccccatt cagtaagttg tactgggata actagctagc 2340
catatcgaga ggattgaaac tgaaccactt ccttacacca tatgcaaaaa tcaactcaag 2400
atgattttaa gacttaaatg taaaacccca aactataaaa actctgggaa ataacctagg 2460
caatccact ctggacatg gaacggaaaa agatttctgc acaaaatcc caaaaaaat 2520
tgtaacgaaa gcaaaaattg acaaatggga catgatttaa cagaattacc atttgactca 2580
gcaatccat tattgtttat atacccaag gaactctaaa catctgtca taagacata 2640
tatacaaaa ttgtcacggc agcactatcac acaatcgcaa agtcagggaa tcaaaactaa 2700
tgtccatcag ttgtagaagg gataaagaaa atgtggtggc agggagtgtt ggctcatgtc 2760
tgtaatccca gcaatttggg aggctgaggg gggtggttca cctgaggtca ggagttttag 2820
accagcctgg ccaacatggc gaaactcgt ctccgctaaa aatcagaaa ttacccaggc 2880
gtggtggcga gcacctgtca tccacgtcac ttggaggccc taggcgtgag aatcgttga 2940
acctggaagg tgggtgttgc agtgagccga gatcctgcca ctgcactcca gctcgggcaa 3000
ccaagcgaga ctctgcctta aaaaaaaaaa aaagaaaatg tggcacatat acaccatgga 3060
atactatgca gccataaaaa agaattggat catgtcctgt gcagcaactg ggatggagct 3120
ggagccatt atcctaaatg aactcactca aaccaatacc acatgttctc 3180
acttataagt agaagctaaa cattgagtac acatggatac aaagaaggga accgcagaca 3240
ctggggccta cctgaggtcg gagcatggaa gatcaaaaaa ctactatct 3300
ggtactatgc tttttatctg gatgatgaaa taactgtgac acaaaaccct ggtgacatgc 3360
aatttacct tatagcaagc ctacacatgt gccctgaaac ctaaaaaaa agttaaaga 3420
aaaacgtttg gattatttcc cctcttctga acaagacat tggtttgccc aaggactaca 3480
aataaaccaa cgggaaaaaa gaaaggttcc agttttgtct attaagcttc 3540
tggggcctac agcctggaga acctggagaa tctacaccc acagaacccc gctttgtccc 3600
caaagaataa aaacacctc ctaaaaaaa aaaaaaaa 3638

```

<210> 262
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_017523

<400> 262
 ttgaaaaagg aaaggtacta caaattcaaa agatttcaact ttaaacactg gcattcctgc 60

<210> 263
 <211> 2461
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_018410

<400> 263
 atgctgggta cgtgcgcgc catggagggc gaggacgtgg aagacgacca gctgctgcag 60
 aagctcaggc ccagctgcgc aggcgtcagc agcggtgat agagaagtac 120
 aaccagccct tcgaggacac ccggtggtg caaatggcca cgtgacctc cgagacgcc 180
 cagggatgta gaatttgggg tgaagacta ataaaggaaa gaaacaaagg agagatccag 240
 gactcctcca tgaagccgc ggacaggaca gatggctcgc tcaagctgc agctggggt 300
 cctagctctc cctgcaccgc cagctcctg ggagccgatt caaaaagcgg tgaggtcgat 360
 ccacgctcag accaggaaga gtcagttgct tgggcttag cactcgagt gctcaaaagc 420
 cctttgaaaa atgaattaa aaggaatac ttgaccacag tggatatact gctcaagggt 480
 gcagagtatt ttgagtgtc aggtaacaga gctggaagg atgtacgtgt gctcgctg 540
 ccttcactgg cctcacctgc cgtgcctgcc cccggatact cgactcgat ctcggaaaag 600
 agtctcgtgt accacgcgaa accagcttca tctccagag aatgggatac ttgcatcct 660
 tcttccacag acattggcctt agtacctaga aatgacagc tctccctaca agagaccagt 720
 agcagcagct tcttaagcag ccagcccttt gaagatgat acatttgcaa ttgaccactg 780
 agtgacctgt acgagggat gctgcactcc atgagccggc tgttgaggac aaagccatca 840
 agcatcatct ccacaaaaac gttcatcatg caaaactgga actgcaggag gaggcacaga 900
 tataagagca ggatgaacaa aacatattgc aaaggagcca gacgttctca gaggagctcc 960
 aaggagacca tcaatccctg ctctgagcct gtgaaaggga caggggcatt aagagattgc 1020
 aagacagtat tagatgttct ttgccgtaa acaggtttaa aattggaaaa agcttttctt 1080
 gaagtcaaca gaccccaaat ccataagtta gatccaaagt ggaaggagcg caaagtgcga 1140
 ccctcgaaat attcttccctt gatttacttc gactccagt caacataaa tcttgatgag 1200
 gaaatagat ttaggacatt aaaaatgttga atttctcgt taaaataagt ttccagacca 1260
 acaatcacag agggcctatg agagaaccgt cagagggaga ttgaaatccg atttgatcag 1320
 ctctcatcgc aatatgtcct gagtccagc aaaccagctc gccgatgtg cctcccgac 1380
 tctcgggcca tgaacagtta cagagggggt cctgcagtc ctggtgctc tcagggtcta 1440
 gaaacccgca ggctgagttt accttccagc aaagcaaaag caaaaagtgt aagtgagct 1500
 ttgtgaaaac taggcacaaag atctctggaa cgaggtaggt gctgcgccaa gcagcattca 1560
 tcttcatcac ttccaaaagc caaccccaca cacagcgcaa ctgcgccga gcagacatct 1620
 gactctcagc ttacgggaaa tagttctgga atattttgaa agtcagtgct acccagcaaa 1680
 actccttcag tcccaagata agaagtgcca ggccacggaa ggaatcgta cगतgaaatt 1740
 aaagaagaat ttgacaagct tcatcaaaag tatgtccca aatctctcgg gcagatgaca 1800
 gtgcctttat tctattggagt gtctacagat tggaaagtcc ggaagtctg atatcaaaa 1860
 gaaggctctc taggaaaaat aaatccagac cctcaactcc aggggtttca gaagtgcga 1920
 tcatcaccct tggggtgcag aaaaagtcta ctgggtccta ctgcaattga ggctcctca 1980
 tctacatgtt ttgctcgtgc catcacaggg gatggcacga gggacatca gttccctgca 2040
 aaaaaccca ggctatcaga accccagggc tccggacgcc agggcaattc cctgggtgcc 2100
 tcagatgggg tggacaacac cgtcagaccg ggagaccagg gcagctctc acagcccaac 2160
 tcagagagga gaggagagaa cacgtcttac aggtatgga agaaaagtga tttcatgcta 2220
 gaaaaattgc aaactaaaag tgtgtagcta ggttatttcg gagtgttatt tatctccca 2280
 ctgtctctct gttttattt ttgtttgttt ttgtattct gagactgtga ggacttggt 2340
 gactctctct cccttaaaat aaatatagtt gaaattggtt ccatacaga taacctcgag 2400
 tcttgggtgt agaaattatg tgaataaaat tgctcaatta gaaaaaaa 2460
 a 2461

<210> 264

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018410

<400> 264

agtatttca tcgtagaaaa attggaaact aaaagtgtgt agctaggtta ttctggagt 60

<210> 265

<211> 1405

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018455

<400> 265

```

caccctcgctc  gcagcctccc  cagcgcagca  gcccggtgt  gggcctgcgg  cagccgggtc  60
tctctggctc  ccacetcctg  gggccgacgg  gccgcaggaa  ggggctcgcc  gggacgcgcc  120
gtcagggacc  tgaggaggaa  caacggaaog  cgttcggaa  ggctcgga  ccgagactc  180
accgcagctcg  gtcggcacacc  gggagaactg  aagcggcagt  agccggcgga  gacgcccgac  240
ccgaaggcgg  gctgcctaggg  agcagacagc  tgaaccgctt  gccagacgct  gaaaccagct  300
gacgcccccc  accgctccac  cgtgctcccg  gctccccgcc  ccgcgcgcgc  gcgggcccc  360
aggcgcatcg  gccgctctgc  ctggaggggc  ccatttcogt  cgtcggtggg  gggaggcaca  420
gtgagtcac  tggggcacgg  cagcgtctaa  gccacaagcc  gagcacataa  gccaggtcct  480
aacggagcct  atgtgtaagt  ccactactgg  tgcagggttg  cacacttcta  agaagagcgg  540
cgtggggggc  toggcgacct  togttcagct  cgctcccccg  tcagtcctcc  tgtgccaaag  600
acacagcctg  atgcttgtgc  tccggtgggc  ggagcttgga  ggcgcgggga  actgcaattg  660
gtggccttga  agcgcgcgcg  agcgggaaca  gctcttgagg  agtgagactg  caggagatgt  720
gggcggtgcc  aaagagatgg  atgagactgt  tgctgagttc  atcaagagga  ccactctgaa  780
aatccccatg  aatgaactga  caacaactct  gaaggcctgg  gattttttgt  ctgaaaaatc  840
actgcagact  gtaaatcttc  gacagagaaa  ggaatctgta  gttcagcact  tgatccatct  900
gtgtgaggaa  aagcgtgcga  gtatcagtga  tgctgcctg  ttgacatca  tttatatgca  960
atctcatcag  caccagaaag  tttggagtgt  ttttcagagt  agtaaggac  caggtaaga  1020
tgttcacctt  ttgtatctga  aacaatttaa  aaattcgctc  aagaaaaatc  ttcagagtg  1080
attaaaaatt  gtgacagtca  gcttcagaga  aactgaggag  aatgcagctc  ggatcgaaat  1140
tgcttgggga  acacagatca  caaagccaaa  ccagtcacaa  cctacctaga  tgggtgacta  1200
ctccacagct  ccgtacgcct  tcaactctcc  ctccatgctg  aggcgcgaata  caccgctctt  1260
gggtcaggag  ttagaagcta  ctgggaaaaa  ctacctcga  caagaggaga  ctattttaga  1320
tattaccgaa  atgaagaaag  cttgcaatta  gtgaacatga  aaggaaaata  aaaattcttc  1380
acagtcaaaa  aaaaaaaaaa  aaaa  1405

```

<210> 266

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018455

<400> 266

```

ccgacaagag  gagatcattt  tagatattac  cgaatgaag  aaagcttgca  attagtgaac  60

```

<210> 267

<211> 927

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018465

<400> 267

```

ggcagcgggc  gaaaggagcc  ggggcctgga  ggtttgcgta  ccggtcgctt  ggtccggcca  60
ccagctgcgc  ccagctgtggt  ttcccataag  gaagctcttc  ttctcgctgt  gcttccacct  120
tlaacccttc  caoctggggag  cgtcctctaa  cacattcaga  ctacaagctc  agaccocagg  180
gagcaaggcc  cagaaagagg  tcaaaatggg  gtttatattt  tcaaaatcta  tgaatgaaag  240
catgaaaaat  caaaaggagt  tcatgcttat  gaatgctcga  cttcagctgg  aaaggcagct  300
catcatcgag  agtgaatatga  gggaaagaca  aatggccatg  cggattcggt  ggtctcgga  360
attctcaaaa  tcttttggaa  ctttttttgg  ccttgacagc  atctctttaa  cagctaggagc  420
gattaaaaaa  aagaagccag  ccttctcggt  ccgattgtt  ccattaaagt  tttctctcac  480
ctaccagtat  gacttgggct  atggaacctt  tttagaaga  atgaaaggtg  aagctgagga  540
catactggaa  acagaaaaga  gtaaatgca  gctgccaaga  ggaatgatca  cttttgaaag  600
cattgaaaaa  gccagaaagg  aacagagtga  attcttcata  gacaaatgaa  atcatgctta  660

```

```

ccaatcaaat ctcaaagcac agaattattg acttgaatca tgggtttttac agttttttaa 720
atgctcaaga ttttgatatt atagatttta ttttaaaata ttaaaatgca agatagtttt 780
gagctatttt aaaaataaaa ttataacatt caacacaaaa tcatggaggt gctctaaata 840
acttttagat ttctctcttc tgtgtgcatt accaatatct aagtgtaaaa ttaataaatt 900
gttttgaatt cctggaaaaa aaaaaa 927

```

<210> 268

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018465

<400> 268

```

ggaacagagt agattcttca tagacaaatg aaatcatgct taccaatcaa atctcaaagc 60

```

<210> 269

<211> 1047

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018487

<400> 269

```

cccactcttc cagccagcgc cccagccctc ccgccgccgc ctgcaggtc ccgaggagcg 60
cagactgtgt cctcgacaat gggaacagcc gacagtgat agatggcccc ggaggcccca 120
cagcacaccc acatcgatgt gcacatccac caggagtctg cctcggccaa gctcctgctc 180
acctgctgtg ctgcgctgcg gccccgggcc acccaggcca ggggcagcag ccggctgctg 240
gtggcctcgt gggatgatgca gatcgtgctg gggatcttga gtgcagtctc agggagattt 300
ttctacatcc gcgactacac cctcctcgtc acctcgggag ctgccatctg gacaggggct 360
gtggctgtgc tggctggagc tgctgccttc atttacgaga aacgggggtg tacatactgg 420
gccctgctga ggaactctgt aacgctggca gctttctcca cagccatcgc tgccctcaaa 480
ctttggaatg aagattttccg atatggctac tcttattaca acagtgcctg ccgcatctcc 540
agctcgaagt actggaacac tccagccccc actcagagtc cagaagaagt cagaaggcta 600
cacctatgta cctccttcac ggacatgctg aaggccttgt tcagaaccct tcaggccatg 660
ctcttggttg tctgattctc gctgctctct gcattcttga cccctctgtg gctgtactgc 720
tggagaatgt tcccaaccaa agggaaaaga gaccagaagg aaatgttgga agtgagtgga 780
atctagccat gcctctcctg attattagtg cctggtgctt ctgcaccggg cgtccctgca 840
cttgactgct ggaagaagaa ccagactgag gaaaagaggc tcttcaacag cccagttat 900
ctgggcccca tgaccgtggc cacagccctg ctccagcagc acttgcccat tctttacacc 960
ccttcccacat cctgctccgc ttcatgtccc cctctgagta gtcattgtat aataaactct 1020
catgttattg ttcccaggaa aaaaaa 1047

```

<210> 270

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018487

<400> 270

```

aaccaaaggg aaaagagacc agaaggaaat gttggaagtg agtggaaatct agccatgcct 60

```

<210> 271

<211> 2280

<212> DNA

<213> Homo sapiens

<300>

<308> U17077

<400> 271

```

ccgccccgca ccagctacgc cccgtccgac gtgccctcgg ggggtcgcgct gttctccacc 60
atccctttcg ccttcttctc gcccgagctg atatttgggt tcttgggtctg gaccatggta 120
ccgcgcaccc acatagtata ccccttgctg caagatgggt tgatgtatgt ctcgctcacc 180
tcgtttctca tctccttgat gttcctgttg tcttacttgt ttggatttta caaaagattt 240
gaatcctgga gaggttctgga cagcctgtac caccgggacca ctggcatctc gtacatgagtc 300
gctgcgctcc tacaagtaca tgccacgatt gtttctgaga aactgctgga cccaagaatt 360
tactacatta attcggcagc ctggttcttc gccctcatcg ccacgctgct ctacattctc 420
catgccttca gcatctatta ccatgatgc acaggcgcca ggccaaagggg gaaatgctct 480
ttgaaagctc caattattgg tcccaaaaag cagcttccaa cgtttgcatc ctggatgaca 540
aacggagaat ccactaaaac gtccacggga ttaacagaa cgtcttgcat gctgagagat 600
gaacacacac tttgttttga catttaatt cactctgctg aataggagga agcttttctt 660
tttctgggga aaacaactgt ctcttggaat tctctgacca tgaacttgct ctctagaca 720
actcacatac aagccctcac tccactaatg gagaatccta gcccaactaa tgccaaagtc 780
gtttggggat ttgtcctcag ctatgggctt ccctagagta ggtctagggg aatactcagt 840
ctgatctttt ttttgtttgt tttattttgt tttttttgag acggagctctc gctcttctct 900
caaggctgga gtgcagtgac gogatctcca ctactgcag gctccgctc ccgggttccc 960
gcatctctcc tgctcagccc tcccgagtag ccgggactac aggcgcgccac caccatgccc 1020
ggctaattta ttgtgatttt tagtagagat ggggtttcac cgtattagcc aggatggtct 1080
gcatctcctg acctcgtgat ccgcccgcct cggctcccca aagtgtctgg attacaggcg 1140
tgagccacgc tgcgcggcct gattctotta aaattgaaga ggtgctgcca aggccttcag 1200
atctaaocga gatgcataga ccttgttctc ggtacttgtt ggcctgtgc tggggagccg 1260
tggtcccgag ttctctggga ggctgacagg gtoaagccac cctgccccac accctccac 1320
tccccctccc ctttctctct cagcattagg attcaaggga aatctgcatg aagccaattt 1380
tgagggtaga cgtgtgggga aaataaatca ttatcacga agacctgggg ctgaggggt 1440
ggggaatggg gagggaaggg catagcctgc tctccatga gctgcacac tcggaaactg 1500
agcagctgcc ggaacgcctg gtcaggaaac acattacac cctcttaagg actggttctc 1560
cagaaagcac cctcagggaa aaaggtgaaa acattacac ctggtattct cctgccacaa 1620
ccgcattgga agaaaaggct gccgcaacat ctacagcagg agtgaaggac ccatgtccca 1680
ggaaacgcgc tgogccacct gcactcaccc cctcacatt ctcttaagca ccgggtggcc 1740
ctccaggctc ggcggaatgg tgggtccccc ggggttgggc aagggtctac caggacctca 1800
acgggcacaa ttgtgcacac taaaatatca aatcaagggt ctggttttta agtaaatgt 1860
ttttctaaag aaagctgtgt tcttctgttg acccagacga atagggcaca gccctgtaac 1920
tgacatgcgc ttctgtcatt gggaatgaaa taattattta cgagaaaagg actgttccca 1980
actggtttga ggcccttacag tttgttatct acattttcc cctctcgggg ttctcgggga 2040
cagggacaga actacaggag tcatgggaaa gaaaattctg gcttcactac tgctcaactg 2100
tcactttctg atcactctga taactttttt ttttttttt ttttgcaacc tgataccttg 2160
aaaaagctct atgtgtctct ccttttgttg cctggcagct gctaggatgt atactgatt 2220
actatttact aagtaccac atgcaataaa aagttgtttg gtaaaatgga aaaaaaaa 2280

```

<210> 272

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> U17077

<400> 272

```

tcaccaggac ctcaacgggc aaagtgtgac acactaaaa atcaaatcaa ggtgcttggt 60

```

<210> 273

<211> 2554

<212> DNA

<213> Homo sapiens

<300>

<308> X87949

<400> 273

```

aggtcgacgc cgcccaagac agcacagaca gattgacctt ttgggggtgt tcgcgagtg 60
gagaggagac cgcccgccgc tgtatttcta gaacctgcct tcgctggcct 120
tgtgacccgc ggcctccctgcc gctgccaagt cggaatattg cgtgtgctcc 180
cctgtggctgc gactgctctgc tgctgcccaa ctggctggca agatgaagct 240
gcgcgcatgc tgcctgtgct cagcgcggcg cgcccgaggg aggaggacaa 300
gtgggacccg ttgttcggcat cgacttgggg accactact cctgcgtcgg cgtgttcaa 360
aacgcgcgcg tggagatcat cgccaacgat cagggcaacc gcatcacgcc gtccatgtc 420
gccttcactc ctgaagggga acgtctgtatt ggccgatgcc ccaagaacca gctcacctcc 480
aaacctgaga acacggtctt tgacgccaa ggcgtcatcg gccgcacgtg gaatgaccoc 540
totgtgcagc aggaacatcaa gtcttgcgc ttcaagggtg ttgaaagaa aactaaacca 600
tacattcaag ttgatattgg aggtggggcaa acaagacat atttgggaaa agaatgtacc 660
gccatgggtc tcactaaaat gaaagaaccc gctgaggctt aacgccaaag aaccaagacc 720
catgcagttg ttactgtacc agcctatatt aatgatgcc acgacccac gcagcttgt 780
gttggaacta ttctgcgcct aaatgttatg aggatcatca acgagcctac gtcagcttgt 840
attgtctatt gcctggataa gagggggggg gagaagaaca tctctgtgtt tgacctgggt 900
ggcggaacct tcgatgtgtc tcttctcacc attgacaatg gtgtcttcga agttgtggcc 960
actaatggag atactcatct ggggtggagaa gactttgacc agcgtgtcat ggaacacttc 1020
atcaaaactgt acaaaaagaa gacgggcaaa gatgtcagga aggacaatag agctgtgcag 1080
aaactccggc gccggtgtag aaaggccaag gccctgtctt ctgacatcaa agcaagaatt 1140
gaaatttgag cctctctaga aggagaagac ttttctgaga cctgactcgg gcccaatttt 1200
gaagagctca acatggatct gtctcggctc actatgaagc ccgtccagaa agtgttgaa 1260
gattctgatt tgaagaagtc tgatatgtat gaaattgttc ttgttgggtg ctgcactcga 1320
attccaaaga ttcagcaact gggttaaagag ttctctaagt gcaaggaaac atcccggtgc 1380
ataaaccagg atgaagctgt agcgtatggt agcgtgtgtc gctgtgtgtc cctctacact tggattgaa 1440
gatcaagata caggttgacct ggtactgctt catgtatgtc cctctacact tggattgaa 1500
actgtggagg gtgttcagac caaactgatt ccaagtaata cagtggtgcc taccagaaga 1560
tctcagatct tttctacagc ttctgataat caaccaactg ttacaatcaa ggtctatgaa 1620
gtgtaagac ccttgacaaa agacaatcat cttctgggta catttgtatc gactggaatt 1680
cctcctgctc ctctggtgggt cccacagatt gaagtcaact ttgagataga tgtgaattgt 1740
attcttcgag tgacagctga agacaagggt acagggaaca aaaataagat cacaatcacc 1800
aatgaccaga atcgctgac acctgaagaa atcgaaagga tgggttaatg tgcgtgagaag 1860
tttctgtagg aagacaaaaa gctcaaggag cgcattgata ctgaaatga gttggaaagc 1920
tatgctatt ctctaaagaa tcagattgga gataaagaaa agctgggagg taaacttcc 1980
tctgaagata aggaagccat ggaaaaagct gtagaagaaa agattgaatg gctggaaagc 2040
caccaagatg ctgacattga agacttcaaa agaaactgga agaaattggt 2100
caaccaatta taagctaaact ctatggaagt gcaggccctc ccccaactgg tgaaggagat 2160
acagcagaaa aagatgagtt gtagaactg atctgctagt gctgtaata ctgttaatac 2220
ggactcagga acttttgtta ggaaaaaaatt gaaagaactt aagtcctgaa tgtaattgga 2280
atcttcacct cagactggag ttgaactgct atagcctaag ccgctgttta ctgcttttca 2340
ttagcagttg ctacatgtc tttgggtggg ggggggaaag aagaattggc catcttaaaa 2400
agcgggtaaa aaacctgggt tagggtgtgt gttcaccttc aaaaatttct atttaacaac 2460
tgggtcatgt gcactcgtgt taggaagttt ttttaccat aagtgacacc aataaatgtt 2520
tgttatttac actggtcaaa aaaaaaaaaa aaaa 2554

```

<210> 274

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> X87949

<400> 274

aactttctc tgaagataag gagaccatgg aaaaagctgt agaagaaaag attgaatggc 60

<210> 275

<211> 1359
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig1632

<400> 275
 tttaagaca gttacctgtt gtgctgctgt tacaatatat aatgaaacca agtcagggga 60
 gtgaattttat caatctttttg atgtaaaagta aaaacgtagt tcacacttca ggagagaact 120
 tcatagcaca atgtctttctc ataagatatt tttaatgatt tagtatttta caacatttgt 180
 ttaccatatt ttgatatacc atttttttct atctgccagag ttttatataa aaaactatat 240
 attattttctc aaagaaacaa tcatattttt atacaaaatt atgttttccag gtaacgaaat 300
 agatgtaggg tcaagtggaa cataagcagt gttaccocgtg gctgggagct agtattatc 360
 acaaaatggt gagctggaac atgcocctgtc tgtgctgtcc ctccctgtctc ggtgtgcgga 420
 tgtgtaggca acattgcctt atcacgctag gttcacctga cactttataa ggaaaaaaag 480
 ttccatagag ttctgtggtc acaaaattgt ttgctttta tcaaatactt taatagaaac 540
 aaagtgcag atattggaat gtaaggaaat atctcagctc ctgcataaga ggattaaagt 600
 atgaaaaggat catttaataga ctgttttact tataagtcac taagtaatcc accatttctt 660
 atggatgatg cttaagcctg gtgaggtttg tactctaagg agcccagatc ataatgcagt 720
 gcaatttctt agcccttaga gtttcttgca aacatttaaa aaaagacata tttaagaaag 780
 aacgataaag aaaaaacata tttaattact gtaaacaggt actgctttat gtttattttc 840
 tctctacttc aaccaaaatc agatctttga ggttttgctg acattgtttg tggttttgca 900
 catgtttctt ctaattggat ttatgaatag ttctatgggt ttccaaagat gaatcatggt 960
 aagaaacact ctgctttttg atccactgtt tgcagcagaa ttatatatc gtatagagaa 1020
 aatccaactt gaataatcca tgttttgat ttggaattg tttttaaaaa taaaagggaa 1080
 aggaaatata taaagctggt atttattctg cattctttac atactatctg ctgttcagta 1140
 taccggtttt ggatatattt gcctctgcac atctacattt gtatatgcaa cagtgcagct 1200
 tatattctaca taaactgtaa ataactcctt ctgtgaaagg atcatcatat caagatgata 1260
 ccaaaagtat gtaaaaaagaa acctgcatta ttttgtaatt atttctata gatatttcat 1320
 ggtaagatta gcagtcataa aagttacttt tttgccttt 1359

<210> 276
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig1632

<400> 276
 gggttttcaa agatgaatca tgctaagaac acttctgctt ttgatccac tgtttgcagc 60

<210> 277
 <211> 994
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig3464

<400> 277
 tgaatgtata tattaagact gtagctgaat tgcacatgaa atcagattgc caactctttg 60
 actttcaatg tttagacattt atccttaagt tgtgagcgat atatgtagca tgcgttgaaa 120
 tgtctgttat agctcttttaa ttcatcagta ttaatacaga attatcattt gcgtttcttg 180
 gtaactttta ttcaatgtaa tcagaagctg tgaatgtttt cctttgtagt cctgtgcttt 240
 tttctgttaa tttttttttt tttttttacg aagcacgtga ctggactaat gtaagcgaga 300
 tgacgtgatc tttaaagctg ctatatatat cagctctcta ctctataagg ttttaaaata 360
 gaataagctt ttatcaaata gataattgat cacttttagg attcacgcaa gtttcagctg 420
 caaatggcgg tctttatagt tcaattctga aaatagcaaa agccaattta 480
 aactgtttct ggcaaacacg acctgtcgtg agatatagtc taaggtagtt aacctatata 540

```

gccttttcaa ctcttaatgc cctccacatg aatcagcagt taagaagggt ctagaaccca 600
tgaagctttt tgtatgtatt actagggttt gtttttctta tgtttgctga ttttacaggt 660
ctgactaaag ctgacctaaa tggatcagtt tatgtgtaat attctagtcg ttaatagact 720
ctttttttct ttggaggagg ggtaacatta ttggacaga tgcagaagga actgtagtg 780
agtcaagaca aacacatctg aaataaagga actgtgtatt aacatgttaa caattcataa 840
ctgcactttt tatgacattt tgaaaatcta ttataggta cagaacaatg gggtttgtta 900
aacgttatca cattttact tgcagaatct tatttcattg ttatttagtg gaattttatt 960
ggttcaataa aattggcaaa actgaacacc aaaa 994

```

<210> 278
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig3464

```

<400> 278
ctgctgtaga tatagtctaa ggtagttaac catataagcc ttttcaactc ttaatgcct 60

```

<210> 279
 <211> 423
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig14683

```

<400> 279
tatgttatgg atatcttatt ttagagtaag aatataaggc atagccatat ttatgaaggt 60
agtaatactc tactaatcaa tacttagaag tttttgttat gactaatctg aatgcttttt 120
agtttttctc taactctagtt atgttggttaa tttataagtc agttttcaga taggaaaga 180
agggtatttg ggggtgtcca ttccactga atagtaagat gatgcttact tagatttcca 240
cagctgtttg aaagctctgt atttggtat aacggaatac ttgttaggg atgcttgatg 300
ttttgtgttt tgtttctaaa ggaagacagt gttttgttcc ttctttagaa aacttgaaga 360
atagaataat gagtccagga ttaatttggg ataaagtctt ttacttcata aattctgatt 420

```

ctg 423

<210> 280
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig14683

```

<400> 280
aggaagacag tgttttgttc cttctttaga aaacttgaag aatagaataa tgagtccagg 60

```

<210> 281
 <211> 391
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig28552

```

<400> 281
atgccattga tgtgaagaag gtgtctgtgg aagactttct tactgacctg aataacttca 60
gaaccacatt catgcaagca ataaaggaga atatcaaaaa aagagaagca gaggaaaaag 120

```



```

aaaaacgtgt cagaatagct aaagaattag cagagcgaga aagactcgaa cgccaacaaa 180
agaaaaagcg tttattagaa atgaagactg agggatgatg gacaggagtg atggataatc 240
tgctggaggc cttgcagctc ggggctgctt tccgcgacag aagaaaaagg acaccgatgc 300
caaaagatgt tcggcagagt ctccagtccaa tgtctcagag gcctgttctg aaagtgttga 360
accatggtta taaccgtgat ttataaaattg c 391

```

```

<210> 282
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig28552

```

```

<400> 282
aagactttct tactgacctg aataacttca gaaccacatt catgcaagca ataaaggaga 60

```

```

<210> 283
<211> 450
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig28947

```

```

<400> 283
ctcatccaag gagctggggc agacttcatt gattctagag agacctgttt cagtgcctac 60
toatccctgc cctctgtgac cagcctcctt accatcacgg ctteactgag gtgtaggtagg 120
gtttttctta aacaggagac agtctctccc ctcttaacct aactctctgg ggtgggaatc 180
agtgcatact gagatggcta gtgtgtgtgt tacgggtttg agttacattt ggcataaaaa 240
caatcttgtt gggaaaaatg tgggggagag gactctctcc tacacgcgca ttgagacaga 300
ttccaaactg ttaatgatat tgtttgttaag aaagagattc tgttggttga ctgcctaaag 360
agaaagggtg gatggccttc agattatacc agcttagcta gcattactaa ccaactgatg 420
gaagctctga aaataaaaga tcttgaaccc 450

```

```

<210> 284
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig28947

```

```

<400> 284
agacagattc caactgggta atgatattgt ttgtaagaaa gagattctgt tgggttgactg 60

```

```

<210> 285
<211> 439
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig30875

```

```

<400> 285
agaaatcaat gacagttgac aggaagagag gacgcataca acaggcaaaa gaggaatgcc 60
cagcagctctt ggtccttgcg gtgcaataact ggccttgagg ccaagtcagc aggggattcg 120
tagtcaactaa cttctaactg aggcagggaa gtaccatggt ctggaaaaag tccaagaaa 180
caggaataga ggcagtgtag caagaggcag atttttggtg ccaaatagat ttgaatcctg 240
gtctctgttc ttcctttgta gagtatgata ttggttcttt cctcccaaaag ctattataaa 300

```

gactaaatgt gtacacaaat ctttgggatg tctgacatat aaatgcttaa caataggat 360
 ttgctggtat tattacaaat gaatttgctt atttttgagc cacttctatg tctgtccatt 420
 aaacccaaat gtgttctgc 439

<210> 286
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig30875

<400> 286
 ggttctttcc tcccaaagct attataaaga ctaaatatgt acacaaatct ttgggatgtc 60

<210> 287
 <211> 338
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31221

<400> 287
 gggaagtac actgcttcac accacaaggc cgtgggaaat cttggaggtt ctgtgccttt 60
 ctgtcacctc tactttttgc agctgtgatt gcactgtccc gcacatgtga ctacaagcat 120
 cactggcaag gaccctttaa atgggtgaaaa tgggcagatg aatagcaata agtggacctt 180
 tgttactctt ctgagttaga aaaattctaa tttagtacac tctgaacaaa gctttattata 240
 ctactctaag atgtgttttg atttgggtgt cagaaaagca cctgacaatg ataatactgt 300
 aactatgata aaattgagaa taaaaagatt ttatttag 338

<210> 288
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31221

<400> 288
 aaatgggcag atgaatagca ataagtggac ctttgttact cttctgagtt agaaaaattc 60

<210> 289
 <211> 417
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31288

<400> 289
 gaatcacttg agccccggag gttgaggctg cagtgaagctg tgtttatacc actgcactcc 60
 agcctgtctg gtaacagagc aagactccat ctcaaaaaga aaagaaaaaa tgctttgcta 120
 cataatgagg ccaggcaaaa aaaaaaaaaa tctctgtgaa atcatataga caaacatttg 180
 caaagctgct actgcatttg taccagtggt aaaaatgtgt ctaccttgca tcttttactg 240
 atttttatga cagattttat attgtaacca tttgagaact ctgtaagtgc tatggcttcc 300
 ttaaaactacg attttatcata tgctccagtg gtttactttg agactgaatg gcaaccagag 360
 aatgtaaaca accaagggtc atctgggttat gttttaaaaa aaagattaat aaaagt 417

<210> 290
 <211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> Contig31288

<400> 290
ggcttcctta aactacgatt tatcatatgc tccagtggt tactttgaga ctgaatggca 60

<210> 291
<211> 394
<212> DNA
<213> Homo sapiens

<300>
<308> Contig31646

<400> 291
gtgctacac cccatgtaaa aagcggaata taaaatgaag attttccagc gcaagatgag 60
gtactggttg cttccacctt ttttggcaat tgtttatttc tgcaccattg tccaagggtca 120
agtggctcca ccacaaagggt taagatataa tgtaatatct catgacagta tacagatttc 180
atggaaggct ccaagaggga aatttggtgg ttacaaactt ctgtgactc caacttcagg 240
tggaataact aaccagctga atctgcagaa cactgcaact aaagcaatta ttcaaggcct 300
tatgccagac cagaattaca cagttcaaat tattgcatac aataaagata aagaaagcaa 360
gccagctcaa ggccaattca gaattaaaga tttta 394

<210> 292
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig31646

<400> 292
gccagaccag aattacacag ttcaaattat tgcatacaat aaagataaag aaagcaagcc 60

<210> 293
<211> 357
<212> DNA
<213> Homo sapiens

<300>
<308> Contig37562

<400> 293
caattatttc aagtgcacct tattaacaaa agtatcagtg gatccaacat aaaattttat 60
agtactaaat gtcaagccta actgtgaatt ttgttctgta tcttaagtaa atttatgata 120
atgttctcga gctatcaaca aaatatatgt acttttggta gctatgaatt ttctaattaa 180
attttacatg ctataaacatg atttttacat gaatgatact ttgtttataa ctatcaaatg 240
tcagtatattt actacaattt tattataaag tgtacattat cactaaatga acttcgattt 300
taaaaatcaa attagcttta gttgtatatt attttttaca aataaagata gacttgt 357

<210> 294
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig37562

<400> 294
atcaaatgtc agtattttac tacaatttta ttataaagtg tacattatca ctaaatgaac 60

<210> 295
<211> 351
<212> DNA
<213> Homo sapiens

<300>
<308> Contig37895

<400> 295
aatagagaca cctctaatta attaaagcgg atgccctccc cactcctccc aggtatttgac 60
tcggagacaca aactcttcac aaacccaaat gtcaggacac catcgccagt gtcactgggc 120
cactctcgtt ggtgtgaggg agccaggagc ccctcagaac tagtaagtct gagaagaggg 180
tgacacggggc ctaggagagg gagaatgag ccggtccaag gtgaattcct tgattctcca 240
ttgtgagtgc accaagaaca agcactccct ccgactgact ctgcctacc aggatctgga 300
acaccttcca ttaatttatt cgttcattca ataatatatt attgactgac t 351

<210> 296
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig37895

<400> 296
ctctgccta ccaggatctg gaacaccttc cattaattta ttcgttcatt caataaatat 60

<210> 297
<211> 418
<212> DNA
<213> Homo sapiens

<300>
<308> Contig38288

<400> 297
gacaagtataa tgggggcccgt tgggacggcg ggtgcctgga gggcagctct gggtcagcg 60
ggcagtgctt agagcacagg cccctctgtt gggggatggg gaggagagca gtcctgccctt 120
gggagcgtag gccccaggga gacttctaaa gccccccctg tcgtctgct ttcaccacag 180
accacagagg cactctgctgc acacacaaagc atctcactgc gccacaggag gggggccaggc 240
ttcctttgcc tgaagctgtt ttgggaagggt tctccacaca ggcaactgac tcccaagctt 300
tggcatgatg gctctttacc atttgataat tttaaacatt gtttttaaac ccaaaacatt 360
tagtggctcg ttgcctctga agatgtaaac aaacaaatcc actatttctg ggaacatt 418

<210> 298
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig38288

<400> 298
tttagtggtc cgttgccctc gaagatgtaa acaaacaaat acactatttc tgggaacatt 60

<210> 299
<211> 413

<212> DNA
<213> Homo sapiens

<300>
<308> Contig38901

<400> 299
taacattttttg tttatgtgtg ggccctgaggt taactgtgac catgggtccag cttgagtggc 60
ttctggagca gccacatttt caaggactgt ccaaaagcca gccagtccag ggctcaggcc 120
tcaccatttg cccactcctg gggagaccat cacctggctc atcggttcca ccaagagtgc 180
ccacagagag tgcccacag acccgctgga ccagcctgct gcgggtcctg gccaggggtc 240
tggttaacgg tgagggtcga ctctgaactg tctctcagtc tccagaaagt gttcaagcct 300
gttgtgttcc caaatctgat tctctctatt gtcttgtaaa tcaaacctcta agtgaanaact 360
tcccatttgt cccctcaag attttttttt attaaatggt tttttaagat cct 413

<210> 300
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig38901

<400> 300
gtttccaaa tctgattcct cctattgtct tgtaaatcaa actctaagtg aaaacttccc 60

<210> 301
<211> 434
<212> DNA
<213> Homo sapiens

<300>
<308> Contig40434

<400> 301
gaatgggtgaa agagagatgc cgtgttttga aagtaagatg atgaaatgaa tttttaattc 60
aagaacatt cagaacata ggaattaaaa cttagagaaa tgatctaatt tccctgttca 120
cacaaacttt acactttaat ctgatgattg gatattttat tttagtgaat catcatcttg 180
ttagctaact ttaaaaaatg gatgtagaat gattaaaggt tggatgatt tttttttaat 240
gtatcagttt gaacctgaat tattgaatta aaatgctgtc tcagtatttt aaaagcaaaa 300
aaggaatgga ggaanaattgc atcttagacc atttttatat gcagtgtaca atttgctggg 360
ctagaaatga gataaagatt atttattttt gtccatatct tgaacttttc tattaataac 420
attttatgaa atcc 434

<210> 302
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig40434

<400> 302
aaggaatgga ggaanaattgc atcttagacc atttttatat gcagtgtaca atttgctggg 60

<210> 303
<211> 391
<212> DNA
<213> Homo sapiens

<300>

<308> Contig40552

```
<400> 303
caccacagccc tgcctcggca cctcgaatcc ctggcgacca tgagtcacca gctccaagcc 60
ttactgtgcc ccagacacaa gagctccatc ccccgccctc tgcagcgttt gcttagcgcc 120
cttgagcgtc cagagccccc tggcccagcc cgtgactcct ctttggggcc tacagatgaa 180
gctggctctg agtgtccctt ccttagaaaag gcttgacctt ccttaccacc cagaaacagg 240
gttttgatgc cctcactagt gttgaagcct gtccagaga gaggtgggac tgcaaggaga 300
ggatggtcag cctcaccac ctgacctgtt tgagcttctt gtttgacaat gtttgcgttt 360
gattttttgt tcaataaaga atttggtaaa a 391
```

<210> 304

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig40552

```
<400> 304
tttgagcttc ctgtttgaca atgtttgctg ttgatttttt gttcaataaa gaatttggtta 60
```

<210> 305

<211> 495

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41413

```
<400> 305
aaatattctt aatagggcta ctttgaatta atctgccttt atgtttggga gaagaaagct 60
gagacattgc atgaaagatg atgagagata aatgttgatc ttttggcccc atttgttaat 120
tgtattcagt atttgaacgt cgtcctgttt gttgttagtt ttcttcatca tttattgtat 180
agacaatttt taaatctctg taatatgata cattttccta tcttttaagt tattgttacc 240
taaagttaat ccagattata tggctcctat atgtgtacaa cattaaaaatg aaaggccttg 300
tcttgcattg tgaggtacag gcggaagttg gaatcaggtt ttaggattct gtctctcatt 360
agctgaataa tgtgaggatt aacttctgcc agctcagacc atttctaat cagttgaaag 420
ggaacaagat atttcagttt caaaattgaa taatgcacaa gtctttaagt attaaaaata 480
aatgtttctt atgtc 495
```

<210> 306

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41413

```
<400> 306
cagctcagac catttcttaa tcagttgaaa gggaaacaag tatttcagtc tcaaaattga 60
```

<210> 307

<211> 409

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41538

<400> 307

```

aaaaaaaaa aaaaaaaaaa aaagagttgt ttctcattgt tcattatagt tcattacagt 60
tacatagtcg gaaggtctta caactaatca ctggttagcaa taaatgcttc aggccccat 120
gatgtcgatt agttctcagt ttctcattcag ttcaacaat aaccaccatt cctgcccctcc 180
ctgccaaggg tcataaatgg tgaactgccta acaacaaaaa ttgcagtcctc attctatttt 240
catccagact tctggaactc aaagattaac ttttgactaa cctcggaata tctcttatct 300
cacttatagc ttcaggcatg tatttatatg tattcttgat agcaatacca taatcaatgt 360
gtattcctga tagtaatgct acaataaatc caaacatttc aactctgtt 409

```

<210> 308

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41538

<400> 308

```

ctcatgttca ttatagttca ttacagttac atagtccgaa ggtcttaca ctaatcactg 60

```

<210> 309

<211> 552

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41887

<400> 309

```

ctgaagacta cgaccatgaa atcacagggc tgcgggtgtc tgtaggtott ctctgggtga 60
aaagtgtcca ggtgaaactt ggagactcct gggagctgaa actgggagcc ttagggtggga 120
ataccaggga agtcacacctg cagccaggcg aatacatcac aaaagtcttt gtgcctctcc 180
aagctttcct ccggggatag gtcatgtaca ccagcaagga cgcctatttc tatttgggga 240
agcttgatgg ccagactctcc tctgcctacc ccagccaaga ggggcaggtg ctggtgggca 300
tctatggcca gtatcaactc cttggcatca agagcattgg ctttgaatgg aattatccac 360
tagaggagcc gaccactgag ccaccagtta atctcacata ctacagaaac tcaccogtgg 420
gtcgctaggg tgggggatgg ggccatccga gctgaggcca tctgtgtggt ggtggctgat 480
ggtactggag taactgagtc gggacgctga atctgaatcc accataaat aaagcttctg 540
cagaatcagt gc 552

```

<210> 310

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41887

<400> 310

```

tactggagta actgagtcgg gacgctgaat ctgaatccac caataaataa agcttctgca 60

```

<210> 311

<211> 745

<212> DNA

<213> Homo sapiens

<300>

<308> Contig42342

<400> 311

```

gcagtaaaga caggagccac ccatgtcaca agaggagcac aggaggggt gttggtgtg 60

```

```

gggcagccct cagggtctcc agaccagccc cactcacac agcagcctag gaaggaaggg 120
cagagtccca ggtgtcagct ggtgggtctc ccaggagctg cccctccctg gaagtccag 180
gacaggaatg acagatcagg gaactgcagg aagctgccac ctctggggct agaatatgcc 240
cagctcgcgg gggctctcta tcggggtctt cgagagccag acagcctgcc ttgtgtgca 300
tacctgcgtt tgctctgtgc agaaccagc acacgtgatt ttgtgtgaca tggcagcagc 360
ctggctccca ggacaggagg cctgccctgg gggaggggct gcagaggagg gggggcagg 420
caccatgag tctgtccagc ctgtcacag atgcatgcc caagctgcgg tcctgatttc 480
agctcacctc agagtaaatc agaataaact gacccagac ttccacgaat gcatgttgac 540
gctttcagtt cacccttttc ttgtctaact ttcttctat ttctttctaa tgcgagagct 600
tattaatcc atatttatca ttttgaataa cttttctcct ttttagtaac aaaaagtact 660
tcactcttag taaaatgtat ttactatttt agtaacaaaa atatacttgc ctaatcatgt 720
ttaaaatata gtgagtga aaatt 745

```

<210> 312
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig42342

<400> 312
caccagaact ttcacgaatg catgttgacg ctttcagttc acccctttct ttgtctaact 60

<210> 313
<211> 398
<212> DNA
<213> Homo sapiens

<300>
<308> Contig43645

```

<400> 313
agttcaaagg cagataaatc tgtaaatat tttatcctat ctaccatttc ttaagaagac 60
attactccaa aataattaaa tttaaggctt tatcaggctc gcatatagaa tctttaaattc 120
taataaagtt tcatgttaat gtcataggat ttttaaaaga gctataggta atttctgtat 180
aatatgtgta tattaaaatg taattgattt cagttgaaag tattttaaag ctgataaata 240
gcattagggt tctttgcaat gtggtatcta gctgtattat tggttttatt tactttaaac 300
attttgaaaa gcttataactg gcagcctaga aaaacaaaca attaatgtat ctttatgtcc 360
ctggcacatg aataaacttt gctgtggttt actaatct 398

```

<210> 314
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig43645

<400> 314
gaaaagctta tactggcagc ctagaaaaac aaacaattaa tgtatcttta tgcctctggc 60

<210> 315
<211> 478
<212> DNA
<213> Homo sapiens

<300>
<308> Contig44289

<400> 315


```

ctaaaaacaa cactcatcag tcttgggaaa tttgaacttt gatcaactta actaaagaag 60
gaagggtagt aagaattttt caaatacaaa tatttgccaa ttcacagatg ataacattta 120
agggcctcaa aagtaagggt ttttccctgt ttctccagtc agcttttgcg aactcctaata 180
gttttttcat aaacattttt tatttgtata attgcaacag ttaagaatat tatcacaaact 240
atttagaaac atttaaaatg ttctttttga tataagctat atacttggaa aaatacattg 300
gtatctaaaa tttgagggtg gttaagactg ctttttgttt taaaaaatgg ttacatttca 360
aatttttgaa gtgttttatg ctccatagtg ctaagtgtga gtttggcaga gttaacagca 420
taagaataaa catgctgtaa ttttaaaaga tgctttgaat aaaaatttat tttaattt 478

```

<210> 316
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig44289

```

<400> 316
catcagtcctt gggaattttg aactttgatc aacttaacta aagaaggaag ggtagtaaga 60

```

<210> 317
 <211> 556
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig44909

```

<400> 317
accatctggg attttctacag cctgggtacc catagccaca coaaggettc tgggagattc 60
tgcaagggtca gctttccagg ctgttcccaa atagctccct gcctcccac tgcccctaaa 120
gccacagcag aagagccatt catctcataa acaaaaaagg agaggaaaga atgaggaagg 180
accctgtgca aggttatatt caggcaggga tgggcttgta cctgacagca cccacccctg 240
tgtggccccc agggccctcat caccctcaga cccctcctaa gcagtccct catgtctctt 300
tggactaggc tgacagcagg aagagcaggg cccatgacgg ggtggaagt cagtttttgt 360
gtctgcttca agaggggggt ttacactctg attccaggac aagcactctg aggcgggtgg 420
gggagagaaa ccttggtctc tcaccagggt ttcacacaca tgtaaatgaa acactatggt 480
agtatctaac acactcctgg atacagaaca caagtcttgg cacatatgtg atggaataaa 540
agtgtttttg aatctt 556

```

<210> 318
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig44909

```

<400> 318
tcaccagggt ttcacacaca tgtaaatgaa acactatggt agtatctaac acactcctgg 60

```

<210> 319
 <211> 710
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig45032

```

<400> 319
aaagataggc ttctaagtta aggcaatca ttcattctgt cattaacaaa atacaaccca 60

```

```

ggcacctgtc atattgccaag tgatattcaa aatggcccat gtagaccttt gtgaagtatg 120
tggcctaaca gacattaaac aaatgtctgt gaaactgaca taataaagta aggtaatgta 180
tatgtgagac attctctttt tataataatt cctgtaaagc agtacttact taggtaatga 240
tatcatactg ttttgtttta ttttttctct aagagctaaa acgtcatcct ctcttcagtg 300
atgtggactg ggaaaactct cagcatcaga ctatgccttt catccccag ccagatgatg 360
aaacagatac cctcatatct gaagccaagga atactgctca gacactgacc gtatctggat 420
ttagtgtgta gcacaaaaat ttctctttta gtctagcctc gtgttataga atgaacttgc 480
ataattatat actccttaat actagattga tctaaagggg aaagatcatt atttaaccta 540
gttcaatgtg cttttaatgt acgttacagc ttccacagag ttaaaaaggt gaaaggaata 600
tagtcagtta tttatcttaa cctcaaaact gtatataaat ctcaaaagct ttttcatct 660
atttttttgg tttattgcac tttatgaaaa ctgaagcctc aataaaatta 710

```

<210> 320

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig45032

<400> 320

```

ttaaactagt tcaatgtgct tttaatgtac gttacagctt tcacagagtt aaaaggctga 60

```

<210> 321

<211> 726

<212> DNA

<213> Homo sapiens

<300>

<308> Contig46218

<400> 321

```

atacatattg ctttagagag caggtaggtg gccatgtgt cagcagtgtg tecttaagaa 60
aataccatct ttctaagcca ctggaaatct tactttacta tttttaacat taatggatgt 120
caggtcatca acctcaagtc tttacatctc catgtatatatt ccataatatat tgtttatata 180
ggcccaagtt tctccttaat tgggatctat atactaccag cacaacatca aaaacatgta 240
attgaatata tcagagctat atatgtaagg aaatgactgt tgacccatt atcatcattg 300
ttgaattcat gttaagtata cctcttaggg gaccataagg caattgagca cataacgaaa 360
aatgatgcac taagaatgta tgcactctct ttgccaaatg catgtgcttt tgtgttaactg 420
ggatgtaaac agaattgcag tgcctgccaa attcttgatc ttggctaaga gagtattttt 480
cccctgtata ttatgactct gagataaaat tgccattttg aaatttcocaa agtaacaact 540
ttttttattt tatgaataaa ctggggattg caatttctct gatctgacaa tcaataactt 600
taacaagat ctaataaagt gtttcaagga aagttttctc aagcaaatgt aatattacct 660
catttgggca tcattactct gtttaattcta tatcaaagga aataaacttg ctacttgcac 720
taaatg 726

```

<210> 322

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig46218

<400> 322

```

accataaggc aattgagcac ataacgaaaa atgatgcaat aagaatgtat gcaactctctt 60

```

<210> 323

<211> 580

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47096

<400> 323

```

gggtgtctct catccttctg tgctgctctg ctaagagatg tccaaggcgg agccggggca 60
agatccttcc agactcatct gtcagagccc caagcccttt agaccagag cccaaggacc 120
atgcttttgg gacattagga ctgcagcctt tgcttctctg tattttggag ttttggtgac 180
tttgtgaccc tggacacact catttggttag ccatagtggg ttcccttggg agcgaacagt 240

gcattgacct ctggatgtca tctgaggtga gaccaccgag gccttttctc tetgtgtaca 300
gaggggaggt agggagtgtc ggactggatg cattacaggg actggggaca gggtagaggg 360
acatccaggg atcaggggcat gagggtgggc aacccccggg cctctgccct ggcatgggtc 420
ccgcatgggg tgagggttag ctgattggct gccacatttc ggccatgctg gctggcgtgc 480
ccatgttgca gatatttccc cgagttcccc agaattggat gtattgaatc tcagccacat 540
gcaacactgt gtccagcatt ctttgcataa aatacttttt 580

```

<210> 324

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47096

<400> 324

```

atattttccc gagttcccca gaatggatgg tattgaatct cagccacatg caacactgtg 60

```

<210> 325

<211> 632

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47563

<400> 325

```

gccatttagt ctgtggtttt ctgttgaagc agtctgaatt gactaaaaa gtcacttgga 60
gtagtataaa accactttcc tgttgaaagc agaacatgct gattcaactg ttttgtcaa 120
tagcaatgat agattttgtt taagtccccc acactttctt atttcaaat gatcaagagt 180
acacttcttg cgagtgaatta agggatgtgt atctaacaga aaaaatata atacctgtg 240
aaccggaata tggaaatcag attgtttctg cccatgtagt catacttaaa aaacaagcat 300
acaaaacaaac atagggaac aaacagcaac cataacaaaa acaaaccttt aaaggtgggt 360
ttttgtcttg ataaatgaat acgggtactct gaaggagaaa aaagtttctc aaatgagctt 420
aaactgcaag tgatttaaaa attagagaat ataattctta aagctattga aagtttcaac 480
cagaaaaacct caagtgaatt ttgtatgtaa atgaaatctt gaatgaagt tctgtgattc 540
tttaagcaaa caattagctg aaaacttggt atgtgttagt ttatgtagat aagtgacttg 600
gcacccatca gaaaataaag ggcattaaat tg 632

```

<210> 326

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47563

<400> 326

```

agcaacaat tagctgaaaa cttggtattg ttgtagttaa ttagtaagt gacttggcac 60

```

<210> 327
 <211> 540
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig48913

<400> 327
 accagagggt gtccttttcc cacagtaatg ggatcggtg gtgtgccttc agggaggaag 60
 agggaggttg tcaagcttga aaaactggct ttagtaggtg tctgactttg ttctccctcc 120
 ccaagtgttc tcaacctcca ttctgcagtg ttcagagttt tagggaaaag gttttgggtgc 180
 cccagcatcc aggtgtttgtg tggcttagcg catgtgaagt gaaaaccttc tggggtgtgt 240
 tgggaagcagc ttctcggttc ttgtgattgt atcctgaggt cccagaacct tattctccca 300
 cgaggatcct cagtgaccat ggtggccaca cgctgtggca tctgtgtggt tctgggtga 360
 gctgaagaac cttgcctgtg gcacttttcc agggtagctg ggaaccgaga gaacatgggtc 420
 cccgtgctgg gactcatgag ggtcatttcc tgcgggctgt gtttcgctgt gctgtgtctt 480
 tatgagcacc atgtaagcct ccttgatttg agataatttg gcattaaaca ttaaaactgca 540

<210> 328
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig48913

<400> 328
 tatgagcacc atgtaagcct ccttgatttg agataatttg gcattaaaca ttaaaactgca 60

<210> 329
 <211> 534
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig49169

<400> 329
 cctaagtgtt acattttttaa aaatacatat ttgggactct tattatcaag gttctacctt 60
 tgttaattta caattcatgt ttcaagacat ttgccaaatg tattaccgat gccctgtgaa 120
 agggggtcac tgggtctcat agactgatat gaagtgcaca tatttatagt gcttagagac 180
 caaactaatg gaagcgagac tatttacagc ttagtatatg tgtacttaag tctatgtgaa 240
 cagagaaatg cctcccgtag tgtttgaaag cgttaaagctg ataatgtaat taacaactgc 300
 tgagagatca aagattcaac ttgccataca cctcaaatc ggagaaacag ttaatttggg 360
 caaacttaca gttctgtttt tgcactctta ttgtcattcc tgtttaatac tcaactgtact 420
 tgtattttgag acaaatagggt gatactgaat ttttacttgt tttctacttt tccattaaaa 480
 cattggcacc tcaatgataa agaaatttaa ggtataaaat taaatgtaaa aatt 534

<210> 330
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig49169

<400> 330
 catacacctc aaattcggag aaacagttaa ttggggcaaa tctacagttc tgtttttgct 60

<210> 331

<211> 602
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig49388

<400> 331
 tgtcagtgga ggggtctctg cagccaaactg agactatctt gctgtgcctt gagccttctt 60
 agggtttaga agaacagcat tcaaaattcc cgtctctgtc agtgtttgcc ttgcacctc 120
 ctccctctaa gcagcgcggtt gggcaaatat gacccacccc ctccctgcag ctccacagg 180
 acgcttctct cctctcccgcc aacccaccca ggctccctgt ggagggtgca gttgtggtac 240
 acgtcccggt tctctgggtt gccgtgaact gggggcggtt cagtcctctt cagccctctg 300
 cctctcccgag tctctgggtt acccgaattt tccacccccc gctctctccc gagggaggtt 360
 agctcttgag caagttggga cttgggcccgt ggcctggaag aatgattggc tgggagggcg 420
 cgggagggag gccagggagg ccggacccagt tgggaggagt gagcaggccc cgggggaggg 480
 ggatgagcgc agtttgcctg ctctctctcc ctgcggcccc cctccgcccc cacacacact 540
 cgggagctct tcattgaaga ttactttaca aaggaaatgt tcaactaata aagaaaaacc 600
 ag 602

<210> 332
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig49388

<400> 332
 cgggagctct tcattgaaga ttactttaca aaggaaatgt tcaactaata aagaaaaacc 60

<210> 333
 <211> 562
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig50728

<400> 333
 gogaatttgg gcccttctgt cctctgatgg gagctgaaag gatgagaggt gggcatctag 60
 atttagggag gctgttcagg ctcttcagg ccttaacctg aacacataga aacctggag 120
 tctgtactgt gtccactgtt gttgtgttgt ctgtgtgtgt tgcgggggat gggcacctgc 180
 atgaatttgg tagagaaaat ggtctctgtc agaggggaaga taagcatagc aaggcaggga 240
 ccagaggaaat cacaggcgcc tggagagcag ccgggcaccc cctccaggga cctgcgggtt 300
 tcctctcagtc ctccaggggc ccagcaactct tcctttaggc cctgtgagcg tcccttctga 360
 ggatacattc tctctatttt ctgaagctga ttgtattggg tgtctgtttc tcgacgcaa 420
 aagagctctg aatgaggaaa gtgcttctgt gctaaacctc cgcgtctctt gaatttcaat 480
 catctatgta ccgcctctga aatttttgcg atatctgtgt accaactgtc cattttactta 540
 ataaagaagt tttctttaaa tt 562

<210> 334
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig50728

<400> 334
 tttgattggg tgtctgtttc tcgcagccaa aagagctctg aatgaggaaa gtgcttctgt 60

<210> 335

<211> 400

<212> DNA

<213> Homo sapiens

<300>

<308> AI497657

<400> 335

tttttttttt	tgcacttatg	gtattttattg	ttggaagatt	gagtaacctta	atgcacacca	60
atgctoaagat	gacttggggg	cacatagggg	actgctgtca	ccatgcctca	ctctgcagg	120
gaaggggctg	ccctactaaa	accccagggg	gcccagtgct	gtgtccagaa	caggctctta	180
tattactgca	gcccacaatg	gaactactga	gtaggagcca	aaagaggagg	gagcaggaag	240
aggtggcatt	tggagagggg	agaccgcacc	cacaggtctg	ccacagcgcg	tcaacgggat	300
gggggtacttt	tacagtcagg	ttgacttcgg	tgccgcacca	ccatctacct	ttgtaggacc	360
actgaaacaa	gggacatcca	ccacggccca	cagccggggc	400		

<210> 336

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AI497657

<400> 336

gagcattggg	gtgcattaag	gtactcaatc	ttccaacaat	aaataccata	agtgcacaaa	60
------------	------------	------------	------------	------------	------------	----

<210> 337

<211> 475

<212> DNA

<213> Homo sapiens

<300>

<308> Contig50950

<400> 337

ctggaagagg	ctcccaaccc	agagtgtccc	tgtgggaggg	aggcagaagg	tgacaattga	60
cacgattttc	tgcacgcgtc	ctcctctacc	ttggaagcag	ttagaatcta	ccaggcacag	120
atgagggccg	ccttgctctga	cggagcttga	tgagcagccc	ttggtctccg	gttccaggag	180
tgagagccca	gctgcctctg	cccacccctc	cccaggccct	tgccagccct	tggtctgaag	240
gtcagggcct	gccccatggc	aggcctgcca	gagcttggct	ggggacccct	ccgcctctct	300
ctccctcgat	gggctggatg	taacttgtgt	ctctagccc	cttaaggagc	ccaggtgttt	360
taaggaaatga	atttgtcact	gcattcttga	tcgattatgg	ttctgagaaa	agcaaatatc	420
acttttggtc	gcattaaaaa	aagcatcata	tataaaaata	agaagatgaa	ggtct	475

<210> 338

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig50950

<400> 338

gtcactgcac	cttgatctga	ttatggttct	gagaaaaagc	aatatcaact	ttggctgcac	60
------------	------------	------------	------------	------------	------------	----

<210> 339

<211> 860

<212> DNA
<213> Homo sapiens

<300>
<308> Contig51660

<400> 339
ggatggcaac cttcagctag actgcctggc tcaagggtgg aagcaatacc aacagagagc 60
atttggtctg ttccggtgtt cctcctgcca gcgaagttgg gcttccgcca agtcgagatt 120
ctgtgccaca cgtactggga gcactggaca tcccagggtc aggtgcgtat gaggctcttt 180
ggccaaaggt gccagaagt ctcctgtgtc caatatgaga tgcctgagtt ctcctcggtat 240
agcaccatga ggattctgag caacctgggt cagcatatac tgaagaaata ctatggaaat 300
ggcatgagga agtctccaga aatgccagta atcctggaag tgtccctgga aggatcccat 360
gacacagcca attgtgaggc atgcactttg ggcatatgtg gacagggctt aaaaagctac 420
atgacaaagc cgtccaatc cctactcccc cacctaaaga ctgggaattc ctcaactgga 480
attggtgctg tgtacctcgc aaaccaagcc aagaaccagt cagatgagcg aaaagaggct 540
aaggggagtg ggtatgagaa attagggcc agtcgagacc cagatccact gaacatctgt 600
gtctttatgt tgcgtcttgt atttattgta gtcaaatgct ttacatcaga atgatgaaaa 660
tagggttgcc acctttctct attttaattc catggtagtc aatgaactgg ctgccacttt 720
aatataactg aaaaatcatt ttgagaccaa gcaggatcaa gttttagaa taaacacttg 780
ttcctagcc atcctctgaa aacagtatga aacatgacca agtacataat ggatttagta 840
ataaatattg tcgaattgct 860

<210> 340
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig51660

<400> 340
gctgcttgta tttattgtag tcaaagtctt tacatcagaa tgatgaaaat aggcttgcca 60

<210> 341
<211> 608
<212> DNA
<213> Homo sapiens

<300>
<308> Contig52490

<400> 341
atcgtggcta gcggacagac acgagcctct tgggaatacc ttgtccatca cgtcatggcc 60
atgggtgcct tcttctccgg catcttttgg agcagctttg tcgggtgggg tgtcttaaca 120
ctactgtgtg aagtacagca catcttctc accattcgca tgatgatgaa aatcagtaat 180
gccacagatc atctcctcta ccgggttaac aagtatgtga acctgttcat gtactttctc 240
ttccgctcgt cccctcaggc ctacctcacc catttctct tgcgttatgt gaaccagagg 300
acctgggca ccttctgtct ggggtatcct ctcactgtcg acgtgatgat cataactcac 360
ttttcccgcc tctctcgtct tgactctcgc cctgagcatg tcccacaaga gcaacacaaa 420
gcaagttct tgactgagaa ctgagtggag gccacagagc ctggggacac aaaaacggag 480
aaggccagaa acagcttcat atggacactg ggacttagcc ccaagcctgt gtgtcctctg 540
aggccagcct ctccaccttc tgagcctcgc cccacactat tgaaaacact aatgaaagta 600
ctcctctg 608

<210> 342
<211> 60
<212> DNA
<213> Homo sapiens

<300>

<308> Contig52490

<400> 342
 ccaggatccat ctctcttacc gggttaacaa gtatgtgaac ctggtcatgt actttctctt 60

<210> 343
 <211> 1282
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig53598

<400> 343
 catgccagca cctttgaacc ggtctcttag aagaagacac acatcctggg tgtacagtgg 60
 tgaatgggg agtgggtgcc cattctgaaa aacgaggcat tcctgctcat tcctctgct 120
 tagctgggtg gcaggggaga gagggaaatg ccaaaaactt ggagtgaagg atgatgctat 180
 tttttatttt taaatatact ttcaggttat ttctctactg ttgcttcaga tctaagttaa 240
 aaggcagatg tccctctctc tccacccccc acgctgaccc cggcctcagt caaggctctt 300
 tgcattgatca cagtctctgt ttctggcctg tggcagggcc gggaagggcc gctggcttcc 360
 gaacagacgt ggttctctc cagcaggcgc atggggagcc cgcgggccct aagctttgtc 420
 gcagatgtca tcattggcag aattacttgt cttgaaaaat aagtagcatt gctgaaacac 480
 acaaccgaat tctctacgat ggccatttgc tcattgtctt tcctctgtgt gtagtgaagt 540
 accctggcag tgtttgctg ctccagatgg cccctcagaa caacagggct ggcttggaa 600
 aaaccccaaa acaggactgt ggtgacaact ctggtcaggt gtgatttgac atgagggccg 660
 gaggcgggtg ctgacggcag gactggagag gctgcgtgcc cggcactggc agcagggtct 720
 gtgtgtcccc caggcagatc tgggcacttt cccaacccag gtttatgcgt ctccaggaa 780
 gctctgggtg cagagtgggt ggccagatctg accatcccca cagaccagaa caaaggaa 840
 tctgggatta cccagtcccc ctccaaccca gttgatgtaa ccactcatt ttttacaatt 900
 acagaatcta tctactcag gctatgggcc tgcctctcac tcagtattg cgagtgttgc 960
 ttgtccgatg ctccggggccc cactgtgtct ctgtgtctta gatcatgggt actccccgcg 1020
 cctgtgggtg gaatcgatgc acggatttgc atgactgtgt tcagatgtgt ttccaacaa 1080
 ccctgtctgt gccctttaat gggatgtgaa gcacttttac cacatggaga aatatatttt 1140
 taatttgtga tgccttttcta caaggtccac tatttctgag tttaatgtgt ttccaacact 1200
 taaggagact ctaatgaaag ctgatgaatt ttctttctg tccaacaag taaaataaaa 1260
 ataaaagtct atttagatgt tg 1282

<210> 344
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig53598

<400> 344
 ccactatttc tgagttaaat gtgtttccaa cacttaagga gactctaagt aaagctgatg 60

<210> 345
 <211> 601
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig53641

<400> 345
 tggaggctgt ggatgatgct ttcaagacaa tggatgtgga tatggccgag gaacatgcca 60
 gggccacgat gagggccag atgaatatcg gggatgaagc gctgattgga cgggtggagct 120
 gggatgacat acaagtcgag ctctgacct gggatgagga cggagatttt ggcgatgct 180
 gggccagat cccctttgt ttctgggcca gataccatca gtacattctg aatgcaaac 240


```

gtgccaacag gaggggccacg tggagagctg gcgtcagcag tggcaccaat ggaggggcca 300
gcaccagcgt cctagatggc cccagcacca gctccaccat cgggaccaga aatgctgcc 360
gagctggcgc cagctctctt tectggatcc agcaccgttg acgaactgca gcatcttac 420
tggccaagcc agagcgcctc ctctcagatt cttctctgac acagcacctt aggcggcttc 480
ttcctgtcag tcggaggtgg catgcaagat gaagctctct ttgctcttcc tgctttcatt 540
ttgtgtttt cctgtgtttt tcatgttttg ggtatcagt ttacattaaa gttgcaaaat 600
t 601

```

<210> 346

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig53641

<400> 346

ctttcatttt gtgcttttcc ttgtgttttc atgttttggg tatcagtgtt acattaaagt 60

<210> 347

<211> 751

<212> DNA

<213> Homo sapiens

<300>

<308> Contig54242

<400> 347

```

aattactcaa agaaggagcc atttcagtta actcaagtga atgaaagact ttgggaatct 60
gcagtggttc cttccctgtt gaccatttgg taacttgtaa tctgacaaa aactcttgag 120
ctgcaacagg ccttgccaga gggctcagga tgggaaaagga agaagggagt aggaaaagaa 180
gaggttaatt tacatttccc ctttaaagta aatttttagcc aactcatcat tctgaaatgt 240
ccctataaag aatgagtcga actagaccag aagccagcct actccttctt acatagcttc 300
tccaacaggg gttagcaatga cctgtccact tcaaacacag ataaggcctg ccatcctcat 360
tggttaaagg cacacgtgag actttcagtg ggctctgctg agaaggaagg cagccocagga 420
gtcaggtatg caggcattgc attgtcagtg tctgcttcca gagtttacac attcaattgc 480
ttccaagggt gaatctcctg ctctgtggaat gctatcagac cccaagggcc aaccttgggc 540
tgggtctatg taogttcttc cgaagcactg atgatcaaaa ttgaagacac attocagagt 600
ttgatttggt gagattaact ggtgtggtgg ttggtgtatg tatgtttat ttttatgtct 660
ttgtatgtag ttctacataa tgcaaatgtg gctttctgat ggacaagacc tcataactgt 720
gattaatatc aataaaaagg ggatgttgtg g 751

```

<210> 348

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig54242

<400> 348

gtaaatttta gccaaactcat cattctgaaa tgtccctata aagaatgagt cgaactagac 60

<210> 349

<211> 637

<212> DNA

<213> Homo sapiens

<300>

<308> Contig54661

<400> 349
 ggcagtgatg tctatgttga gattaactta tgtattgagg aaaatttgaa gtttattttt 60
 tcgatgaata aggctgtcaa atgatttagt atagattaat gacatctttt ttagaatat 120
 taaagtgatg attcctcatt atgtcatcat ttctgataat tagatgtcta atttgaatgt 180
 tagataatgt ttccacatct atacctattt ctttctaggg caoctctgac cctggggcctt 240
 ggggtaggcc tttagggcac agtagtgtct gtgttaagtt cactaaatgt gtatttaagt 300
 agaaacattc ctatgtaaaa atgtgtgtat gtgaacgtat gcatacattt ttattgtgca 360
 cotgtacatt ctgaagaaat agtttggaat ttgttaagc acaaaccata aaagagtgtg 420
 gagttattaa atgatgtagc acaaatgtaa tgttttagct ataaaaggct ctttctattt 480
 tctatggcaa agactttgac aottgaaaaa taaaaccaat atttgattta tttttgtaag 540
 tatttaggat attatttaa ataatgatt gtccattatc aatataatag ttgtgaaatg 600
 atttaagtta ataaactta tgcttctgtg tctgttg 637

<210> 350

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig54661

<400> 350

ctgtacattg tgaagaagta gtttggaaat ttgtaaagca caaacccataa aagagtgtgg 60

<210> 351

<211> 924

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55188

<400> 351
 gcgacaagta ccgcaagcgg gcaactcatc tgggtgcact gctggccctt gcgggcctct 60
 tgcgcgcctt cgtgtgtgtg ctgtacatct acccatttaa ctggcccttg atcgagcacc 120
 tcaactgctt ccccttcacc agccgcttct gcgagaagta tgagctggac cagggtgtgc 180
 actgaccgct gggccacacg gctgcccctc agccctgctg gaacagggtc tgccctgcgag 240
 ggctgcccct tcgacagcgc tctctgtgtg ccagagagcc agagacccaa gacaggggccc 300
 gggctctgga cctgggtgcc cccctgccag gcgaggctga ctccgcgtga gatggttgtg 360
 taaggcgggg tttttctggg gcgtgaggcc tgtgagatcc tgacccaagc tcaggcacac 420
 ccaaggcacc tgccctctct agtcttgggt ctcagttct aatatccgc tccttctgta 480
 gaccatctcc tggggcaggg tctttttctt cccaggtctc cagcgtgccc tctgtggtg 540
 ccttctcccc cactactact ggagcgtgcc cttgtgggg acgtggtgtg gccctcagtt 600
 gcccccaggg ctgggtgccc accatgcccc ttctcttttc tctctctacc tctgccctgt 660
 gagcccatcc ataaggctct cagatgggac attgtgggaa aggccttggc catggtctgg 720
 gggcagagaa caagggggga gacacaagta gacctcaggt agaacgacac tggcgaggac 780
 caccocaggg cctgctccca gggagtgtct gaggcgcac aggcccgctt ttaccagtt 840
 tatatcacgg tcttcaattt taaaagtaac gctaacttgg tacggacgat gtctcatgta 900
 ttaataataa ttctttatgg cagt 924

<210> 352

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55188

<400> 352

agtaacgcta actttgtacg gacgatgtct catggattaa ataattttct ttatggcagt 60

<210> 353
 <211> 699
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig55353

<400> 353
 tgattatgcc aagagctcta aacagaagtt tgagaaggtt aaaattaagt tgtagtatct 60
 gagttgtttt tattttcttc ctttgggtgt tatgaaggtt ttcataagaa cttaatttc 120
 aggggaaaaa atgcctgatt tgctattttt gacatttctt cgtctcttaa gaagtcagtt 180
 aaatatgttt tcatagttta tattcctgtt tcatagatta ctgtgaaaaa tgtattttaa 240
 cctatgaatt ataaaaagtt atttagattc tagcgtgagt taaatagatt agtcatatat 300
 cttttagatt tgtggatttg acatgtaaat tatgtgtgtt gtataagtaa gttagtact 360
 aaacatatgt catggttatt gataaacttg ttgctatttt tttccaaatg ctatcagttg 420
 tgtggacttt ttaaaaaatta gtttgaattt tgggaatgttc tgtgataaaa tataatttca 480
 actattttgt acattttaat atgccatgtt gtatatgtct gtatttaaaa atgttgttaa 540
 tatctgcatt ttaagaatta tgaagatttt tcttcaaaaa tgacagaact ctccatactt 600
 aattgtgaca cattataaga tatctgattt taagcttttg gattttgttc taaaaattaa 660
 gtttaaacat gctgaaaatt ccataaaaaa aaaattttg 699

<210> 354
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig55353

<400> 354
 taaaatagta tttagattct agcgtgagtt aaatagatta gtcatatata ttttagattt 60

<210> 355
 <211> 809
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56503

<400> 355
 gcatgtgaga tgagtactg cgggtgaatg tgtccacagt tgagaggttg gagcaggatg 60
 aggaatcctt gtcaccatca ataactactt gtggagcgcc actctgcccc agacgccacc 120
 tggcgagaca gcatggagct ctccatggcc aggcgtgctg tgtgcatgtt cccgtctctg 180
 tgcccctttg ccgcctctct gcaaaacctca cagggtcccc acacaaacgt gccctcaga 240
 agcagccctt cggaggcaga ggaaggaaaa tggggatggc tggggctctc tccatcctcc 300
 tttctcctt gccctgcgat ggctggcctt cccctccaaa acctccattc cctgtctgcc 360
 agcccttttg ccatagcctg attttgggga ggagggaagg gcgatttgag ggagaagggg 420
 agaaagctta tggctgggtc tggtttcttc ccttccaga ggtcttact gttccagggt 480
 ggccccaggg caggcagggg ccacactatg cctgcgcctt ggtaaagggt acccctgcc 540
 tttacacaga gccctggcat gttcctgccc cacaggaata gaattggagg agctccagaa 600
 actttccatc ccaaaggcag tctcctgggt tgaagcagac tggatttttg cctcgccct 660
 gaccctgtgt cctcctttga gggaggggag ctatgctagg actccaaact cagggactcg 720

ggtggcctgc gctagcttct tttgatactg aaaactttta aggtgggagg gtggcaaggg 780
 gtgtgcttaa taactcaatt ccaagcctc 809

<210> 356

<211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56503

<400> 356
 gaaaactttt aagggtggag ggtggcaagg gatgtgctta ataatcaat tccaagcctc 60

<210> 357
 <211> 976
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56678

<400> 357
 gaaggatata ctttgttata acttattatt ttgttctctg taaatacaag atgtttatag 60
 gaaatatgta ttctgaactc tatctgcaga atgagtcact acaccaaaat agttctatta 120
 tttagaatgt gttaatttta aagggaacctg ataggtattt atttaccatg gcgatccaca 180
 tttgtgtgaa agcatgtgat catactaacc cagcctcctg gaatgtcgct gtacgatgat 240
 tgatgtcttt ttctcagtcct atagttacaa ttgtttagta tgctaatacg tccagttccc 300
 tgagggtttaa gatcaaatat aaattactct gcttttcgac tcattcaggt agcatgttac 360
 ctgaacctga ttgctacttt ttcatcttaa atattatatt tcctcatcta atctgccttc 420
 cctcatcca cagacatttg gagaaggaaa tgggagggtg tctgttatcc ctctctcttt 480
 gctttgtccc cggtgttaga ctggcagcgt cagttgctcg gtgggcttgg tttagagccgt 540
 gggtgaggca ggtggctggc ggggacaggg agaggctgag agggaggtgg tggcatttac 600
 tgctctgaca ctccactgt cctctgctgg gatgctgggg ccaaggcctg tggggcgtgt 660
 gaactgcaca gccaggagca aggaaccac taaatactcc gtcacctcca gtccocctct 720
 acagtgtaaa attattacat aagcaggtga aaggtagaag gcgaattatg tgagtaataa 780
 tggctgtttt tctcttcagc aaaaatgact attttgtgt gtgactaatt tattttttat 840
 attgtaaga tacaataaac cgggtgaaat atctgctttg ttgacaaagc tgtgctttct 900
 atggccttat tcgctgtctg ttctcctgca aatagcgccc tctaaaaaga agagtcagac 960
 aataaactgg ttgaaa 976

<210> 358
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56678

<400> 358
 tattacataa gcaggtgaaa ggtagaaggc gaattatgtg agtaaatatg gtctgttttc 60

<210> 359
 <211> 1118
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig57584

<400> 359
 agctgttgtg catccagagg tggaattggg gcccggcatt ccctcctcgt cccgggctgg 60
 cacttgcccc caccctgcaa ctctgtgttg agatgggctc agccaagagc gtccagtcga 120
 caccagcgcg gcctcgcgcg acaacaagca tctggctcga gtggcgagcc cccgttcacc 180
 tagtgtggc atctcgcgca ctcccatcca ggtggagagc tctccacagc caggctacc 240

agcaggggag	caactggagg	gtcttaaaaca	tgcccaggac	tcagatcccc	gtctctctac	300
tcttggattt	gcacgggacac	ctatgaagac	cagcagtggg	gaccccccaa	gccactgtgt	360
gaaacagctg	agtgaagtat	ttgaaactga	agactctaaa	tcaaatcttc	ccccagagcc	420
tggtctgccc	ccagaggcac	ctttatcttc	tgaattggac	ttgocctctg	gtacccagtt	480
atctgttgag	gaacagatgc	caccttgga	ccagactgag	ttccctcca	aacaggtgtt	540
ttccaaggag	gaagcaagac	agccccagag	aacccctgtg	gccagccaga	gtctccgaaa	600
gccctcaagg	gacccctgaga	ctcccagatc	ttcaggttct	atgcgcaata	gatggaaacc	660
aaacagcagc	aaggtactag	ggagatcccc	cctcaccatc	ctgcaggatg	acaactcccc	720
tggcacccctg	acactacgac	agggttaagcg	gccttcaccc	ctaagtga	atgttagtga	780
actaaaggaa	ggagccattc	ttggaaactg	acgacttctg	aaaactggag	gacgagcatg	840
ggagcaaggc	caggaccatg	acaaggaaaa	tcagcacttt	cccttggtgg	agagctaggg	900
cctgcatggc	cccagcaatg	cagtcaccca	gggcctgggt	atatctgtgt	cctctcacc	960
cttcttcccc	agggatactg	aggaatggct	tggtttctta	gactcctcct	cagctaccaa	1020
ctgggactc	acagctttat	tgggctttct	ttgtgtctgt	tggttttctt	ttatatataa	1080
ggaagtaatt	ttaaatgtta	ctttaaaaag	gtatatgt	1118		

<210> 360
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig57584

<400> 360
 aggaatggct tggtttctta gactcctcct cagctaccaa actgggactc acagctttat 60

<210> 361
 <211> 859
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig63649

<400> 361
 gtgcagggt accagtgtgc ggagtctctg ttgccaaagt gaaggtggcc ctgggcaggc 60
 acaggtgtgg tcatatcttc agccaacagg accatcctcc ggagggccac ctctggggag 120
 ttccctacggg aagagagtga cagatttggt gcttctgtgt gttctgccc ctccagtgagg 180
 gccgctcgcg gagacagcgg gtggatctct cagcagcctg tctgtctgagc ctgctctctc 240
 aagttactct ttaaaatcag gaccgggtcg tgtccagacc tacaggccct gtctccgctc 300
 cccagggcctg caggagtgtga gggctgcacc tgctcgctgg agagggagag gcagatttag 360
 tggagccctg gcatggactc ggaactggcct ttggaagctc cctgccctca cgggttgccct 420
 gtccaccactg cgaagtgtagg cttggaggac ctgcacactg gaaaggctgt gtgtgttctt 480
 gggtccacac ctgccagagc taacttactg ccagacggcg acttactgtg ggccaccctc 540
 agtgnaaccgg ggtgtctcca gctggcccta cagagcactt ctgtgctggg gatgagttag 600
 aactctggcg gaggggggc ccagcgccgc cctcgatac agccctgctc tgccctctgc 660
 ccgtacttat accaggtggg atccctgcc tgcattgcct ggggatggg tggtgtggg 720
 cagcgcctcg tgtggaactg gatgttttca gggagccagc cctttctcca tgtcaacaca 780
 gttcacaaata tagtttttcaa agtacagttt aaaactcaaa agtaaaacttt tcagcaactc 840
 aaaaaaaaaa aaaaaaaaaa 859

<210> 362
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig63649

<400> 362
cagcctttcc tcatgtcaac acagttcaca atatagtttt caaagtacag tttaaaactc 60

<210> 363
<211> 1170
<212> DNA
<213> Homo sapiens

<300>
<308> Contig63525

<400> 363
gccatggctc cctggcgga gcgagcactc gcgctgaac ccgctgcgcg cgggtgtggct 60
cagcgtgacc gcgcgttcc tgctgaccct actgctgcag ctctgcgcg ccggcgtgct 120
cccggtgtgc gcgatcttcc aggaacctgat ccgctatggg aaaaccaagt gtggggagcc 180
gtcgcgcccc gcgcgtgcc gagcctttga tgtcccaag agatatTTTT cccactttta 240
tatcatctca gtgctgtgga atggcttctc gctttgtgtc ctactcaat ctctgttctc 300
ggagagcact ttccaaagt ggcttcatgg ttgtctcaga attctcgagg cggcacagtt 360
ccagggaggg gagctggcac tgtctgcatt cttagtgcta gtatttctgt cggctgcacag 420
cttaaggaaga ctcttcgagt gcctctacgt cagtgtcttc tccaatgtca tgattcacgt 480
cgtgcagtac tgttttgac ttgtctatta tgtccttgtt ggctaaactg tgcgtgagcca 540
agtgccaatg gatggcagga atgctacata acagggaaaa atctattgat gcaagcacgg 600
tggttccata ttcttgggat gatgatgttc atctggtcat ctgcccatca gtataagttgc 660
catgttattc tcggcaatct caggaaaaat aaagcaggag tggctattca ctgtaacacc 720
aggatcccat ttgggagact gtttgaatat gttttctccc ctaactactt agcagagctg 780
atgatctacg ttcccatggc cgtcaccttt gggttccaca acttaactgt gtggctagt 840
gtgacaaaag tctcttttaa tcaggccctg tctgccttcc tcagccacca attctacaaa 900
agcaaatatt tctcttaccg gaagcatagg aaagctttcc taccattttt gttttaagtt 960
aacctcagtc atgaagcagt caaacaggt gatggtttca atgcctaagg acagtgaagt 1020
ctggagccca aagtacagtt tcagcaaacg tgtttgaaac tctccattcc atttctatac 1080
cccacaagtt ttactcgaat gagcatgcag tgccactcaa gaaatgaat ctccaaagta 1140
tcctcaagaa attaattact aatggcagat 1170

<210> 364
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig63525

<400> 364
ctottaccgg aagcatagga aagctttcct accatttttg ttttaagtta acctcagtc 60

<210> 365
<211> 632
<212> DNA
<213> Homo sapiens

<300>
<308> Contig64688

<400> 365
aagaatgcta agatgatttc agatatcgaa aagaaaaggc agcgtatgat tgaagtcag 60
gatgaactgc ttcggttaga gccacagctg aaacaactac aaacaaaata tgatgaactt 120
aaagagagaa agtcttccct taggaatgca gcataattct tatctaattt aaacagcgtt 180
tatcaagatt attcagatgt tcaagctcaa gaaccaaacg taaaggaaac gtatgattca 240
tcacgccttc cagctctgtt attttaaagca agaacaacttc tgggagccga aagccatctg 300
cgaaatatca accatcagtt agagaagctc cttgaccagg gatgagaaga gcagcttact 360
aaaatgtgcc tataggaaga ctagtctcat gctgttaact tctgaaactg tacctttata 420
aatcaattgt tttgcaagaa agttatggcc tacttagaat ctaaaatttg ttattcaaat 480

taaatggctg tgaacaatg taaatagcat cagtttgc aatagtntta aaggccataa 540
 tcatcttttc tggtaaatat cttagtagat tttaaaatgt tgacacctta atcggtccca 600
 ggtatgagcc ataataaact tgtaaaatta ag 632

<210> 366
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig64688

<400> 366
 ggcgtggaac aatgttaaat agcatcagtt tgtccaatag ttttaaaggc cataatcatc 60

<210> 367
 <211> 2596
 <212> DNA
 <213> Homo Sapiens

<300>
 <308> AL080059

<400> 367
 gatgtcaact gctttatttc tggctttggc ctgtggtctg tgatacccat cctgcttgat 60
 gttctgcaga atggcacttg actgctgggc atgcatgaag ttaagggcga gaaacagtat 120
 gccatgtggt ctgtaccatc atgtgtctct tcttctctct gggccctctc actggtgaa 180
 ttcatcaag atctggccoa tgcggtgtca ctatcaagcc attaatgttt gctctgggtg 240
 ctgtgcagcc cagttggctt cctgggtcaac aaggacctca agaactgcct gtggacagag 300
 gccctctacc agtgcagtga acacacacct accctcccca gctctccagg aacctcatg 360
 gctgccagac tgaatggggg gctggatagt gtggacatgt gttcaactgt attatgctgt 420
 ggcctccaggt gagggtgagg actgggccta tatagaatcc agataccatt gtcaactctc 480
 cttattcccg tctaagatgt gacgagagtg ccatagttag gggtctggga agaggtatct 540
 ctgattttgt ggcctctgct tgcctgactt caggtgcact atactcttta ttttgccttg 600
 ctgcctctat cctcatttcc ctccctctca ttctctcttc ctcctctcct ttccgtgtag 660
 cctcctttcc tccctctctg ccttccctct ccttctttcc ttattctttt ttatttttgt 720
 taaatagtac cacagagaaa acaactgaaa aaccacattt ttctacatac agctggggag 780
 gtatgtgaga acttggcact gcgcacacat actaggttga aagagaggtg agggaaaccag 840
 aaggccaaat ggaatctgtg gcaaaccttg aacctgtctc ctgcgcttgc tctacagttc 900
 tgaagtgtga aatcgttttc atgcctagca tctgcttgag ttataaacc caaggcagcc 960
 atgtcataga ctagtgttta ctctgtttt gactttgttt taatgcttcc taagacccaa 1020
 gtgcctctgt ctgtttcttc ctttgggtga gctctggccc atctggacct caatccccag 1080
 ctttcccaat ttcagcagtc ctttgccttc ttgtcttcta cctcaaatga cccacaggat 1140
 gggctttagt ctccaatag gagcatctca agcttctcct gggggatggg gatgggagt 1200
 ggcggaatct gttttggatc tccgggttat tccagtgagg tgtaaaagca gagctgggac 1260
 ttctccctct ttatccctga ggggtgggtaa gaaggactgt atctacacct gttctctcct 1320
 accctctctt ttgttaggga ggcctcattc taagtctctc aagagagtcc ttgctctaaa 1380
 gctgtagcaa ggtgtgtcta ggtgggggat ttggagcaaa accgtcagat aggcagatga 1440
 ctggtatgga cctgggcctgc aaaatcagac agaaatggct tgagaagcgg cagggggagc 1500
 atgcctgtct ctcaagtata gagtatggga gggacctccc tagcttggaa aatgagaatt 1560
 gaagggttta tgaacaataa ggaatgcctag ttgaggtagt tcccaaatgt ttgtccaatc 1620
 ttatcattga tagattttat aagccacaga gacaaaccag aaacggaata atgttacttt 1680
 ggaatcctta tttttttgtt ctagggtgtg ctttgtacat gcagaagaat gctatatgct 1740
 gcacattttg ccttttaagt cttacagact tcccattttt agtctaatgg gaagatacag 1800
 atgtgcaagt ctgctttttt gtttttttgt attatttttt tttttgctct gtgttatgga 1860
 catgcacaga catgcacaga agtgagagag atggctcttg gaccccatgt gctcatcacc 1920
 tagctgcact acttatcagc tatggtcaac ctggtttcat ctgtatctct cctctttcac 1980
 ctgattttgt ttattgaaat ccaagacact atgccaatgc aacctgtagt actttgggag 2040
 attggtagtc tcttttgatg gtgtagtgta tgggggtcac tatcataatc acatcaggtc 2100
 tgccttttgc ttttaattgt aactaatgaa gttccagaga tgggccttag aaatgtgttt 2160

taagaattaa	caaggagtct	caaaaagaaa	tgagagggat	gcttcccttc	ccttgcact	2220
acaaaacaag	agagagactg	tctctgttga	aaactctctc	aaaaattctg	atatggtaag	2280
gtactctgaga	cccttcacca	gaatgtcaat	ctttttttct	gtgtaacatg	gaactctgtg	2340
tgaccattag	cattgtgtatc	agcttgtaact	ggctctcataa	ctctggtttt	ggaagataaa	2400
tttggaaatt	gttgcctgtg	tctgtgaaaa	taacctcccc	aaaataatta	gtaactgggt	2460
gttctacttg	gtaatttgac	acccctgttaa	taacgcgaatt	atttctgtgt	tctttaaaccg	2520
tataataagt	tgtaaagttg	catgcatgat	ggaaaaataa	aaacctgtat	ctctgtcaaa	2580
aaaaaaaaaa	aaaaaa					2596

<210> 368

<211> 443

<212> DNA

<213> Homo Sapiens

<300>

<308> AA555029

<400> 368

tttttttttt	tttttccatt	aacccaacac	aagtttattt	actaagtcaa	gtcaaaactga	60
ggagtatattg	tctctttggg	agttggtaga	caaagaatac	atatacatat	tctatttccc	120
attaagcatt	cgatcatgtt	acaaaacaat	ggcctagaga	actatcattg	gaactttacc	180
aaatctcgctt	gaagcaaaat	gtgataaact	gcgaaatagg	tggagaatga	cactgggaagt	240
cataaaagttg	attctcctaaac	acgggtttga	actgcatagg	accgcttata	tgcattatttt	300
gataaatata	ttgaaaattt	ttttggaaat	ttgcaacaat	ttaaaaaact	tgacagatgca	360
ccatgtagct	tagaaatact	gaaaaactaa	gaaaaaggtg	tgtcatgaat	tcataataa	420
tatgtagaca	ctagtctatt	tta				443

<210> 369

<211> 3160

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003748

<400> 369

attctgcgcg	cgcccgggcg	gacgatccag	cgaacagccc	cgcttctaac	cagagatgct	60
gctgcggcg	cccgcgctcc	gccgcgccct	gctgtcccg	ccctggaccg	ggcgccggcct	120
gcggtgggaag	cacacctcct	ccctgaagg	ggccaacag	ccctcttag	ccttcaacga	180
gggcagccct	gagcgagatg	ccctgcaaaa	ggccttgaag	gacctgaagg	gccggatgga	240
agccatccca	tgctgtgttg	gggatgagga	gggtgtggac	tcggacgtgc	agtaccgaat	300
gtcgctcttt	aacctggac	ataaggtggc	caagttctgt	tatgcagaca	agagcctgct	360
caacaaagcc	attgaggctg	ccctgtgctg	cgggaagag	tgggacctga	agcctattgc	420
agacggcgcc	cagatcttcc	tgaaggcgcc	agacatgctg	agtgggccgc	cgagggtctga	480
gtactctgcc	aagaccatgg	tgggacaggg	taagaccctg	atccaagcgg	agattgacgc	540
tgcagcgga	ctcatcgact	tcttccggtt	caatgccaa	tatgcggtgg	agctggagg	600
gcagcagccc	atcagcgctg	cccogagcac	caacagccag	gtgtaccggg	gtctggagg	660
ctctgtggcg	gccatctcgc	ccttttaact	cactgcaatc	ggcgccaacc	tgccggggggc	720
accggccctg	atggggcaacg	tggttctatg	gaagcccagt	gaacctgcc	tgtcgccag	780
ctatctgttc	taccgcaccc	tccgggaggg	tggcctgccc	cccaacatca	tccagtttgt	840
gcagctgat	ggggcccatat	ttggggacac	tgtaaccagg	tcagagcaac	tctgtggcat	900
gaacctca	ggcagtgctg	ccaccttcaa	acacctgtg	aagcaggtg	cccagaacct	960
cgacgggttc	cacaccttcc	cacgcctggc	tggagagctg	ggcggaaga	acttcaactt	1020
cgtagcaccg	tcggccgacg	tggagagcgt	ggtagcggg	accctccgct	cagccttcga	1080
tacgggtggc	cagaagtgtt	ccgcctgctc	cgctctctac	gtgccgcact	cgctgtggcc	1140
gcagatcaaa	ggggcgctgc	tggaggagca	cagtcggatc	aaagtggggc	accctgcaga	1200
ggattttggg	acctctctct	ctgcagtgat	tgtatgccaa	cctttgcccc	ctctcaagaa	1260
tggtctggag	cacgcgcgct	cctgcgccag	cctcaccatc	ctggctgggg	gcaagtgtga	1320
tgactccgtg	ggctactttg	tggagccctg	catcgtggag	agcaaggacc	ctcaggagcc	1380
catcatgaag	gaggagatct	tggggcctgt	actgtctgtg	tacgtctacc	cggaacgaca	1440
gtacaaggag	acgctgcagc	tggttgacag	caccaccagc	tatggcctca	cgggggcagt	1500


```

gtttctcccag gataaggagc tctgtcagga ggcacacaa gtgtctgagga atgtctgccg 1560
caactctctac atcaacgaca agtccactgg ctgtagatgt ggccagcagc cctttggggg 1620
ggcccagagcc tctggaacca atgacaagcc aggggggcca cactacatcc tgcgtctggac 1680
gtcgcgcagc gtcatacaagg agacacataa gccctctggg gaactggagct acgcgtacat 1740
gcagttagacc cctctcgggc tccaccgtcc agctgtctgt ccgtccaggt ggccgacctc 1800
actgcacaga ccccactcca gccctctcac ccctctctca tgacacagct cctttctata 1860
atccgggctt gactcccttc ttaccactgt attctggcct ctcccatgcc tcaggctctg 1920
gtttgagatc gtgtctggga ggaacatggc cactaccctc tatcccatgc gccatgtggg 1980
aggtatgacc ctggtgcctg gcaggttctc cctctgccct ccactgggcc cagtggtcca 2040
gggacctggg gaaaggagat ggagcagctc ttgggatctc ttggggaaaa gaggcccat 2100
ctggggccctc tggcaaacct caccactcac agaggctcct ggccctgac cctgcccctc 2160
caggtgtcca gggtaaatg taactcagac tgacctgtgg ggacagggg gcaccagctg 2220
gccctgcctc ctctggtctg ggtgtctcac ctctcctcact gtatctttgc ccagaccac 2280
ctggggcaggt agggccctgt cccagccacc acaccttaga tgctggcatg ccttactcca 2340
ggtgcctgtg tttggccgag gctgtgtgta ttccgggtct gcaccacatg gcgggggttg 2400
ggggcgctg gaggccacct gccaaaggct gggatgggat ggtctcgccg gtttagggcg 2460
tgattctgga aaaccttgga tgggccttcg tcctatgtca gccttccctt tgatcctcag 2520

```

```

gccctacctg tagagacctc cactcctaga gccagtctca ggtgtcggga tttccctgca 2580
ggagctcagc caccactgtg ccatggtgac acagagccaa gcagacatgt gccctccctt 2640
ctccagcccc ccagaggcct ggccctgggt tctgcagcat gggcgagaa cgttgcctgt 2700
agaactctcc tctgcctggg agtggctctg tgtggacacg tccctcactg gccactctt 2760
tltttacgca agcaactctg tgaccacgat tctctccaca gatgcctcct ccttggatc 2820
tgagtgtgca gagatctgta aagcatgact ttcaaggatg gttcttaggg cagtgtgaaa 2880
gtgttggttc tttcctcagg atgcctgcat gggaccccac cggagagctg tgtggccatt 2940
ccccaaagtgc cactggccca tggatggggg tgggtgtctg tgccagctgg gctgggtgtg 3000
gtttctgtgt ccttccagga tatgtgtcat tctccatgag ggtctcgccg aggtggctgg 3060
gtgggggac aggctggagt attcttagtt ctactggttc tacactgtga ggtggcaatg 3120
ggatttcttc agatgccacc caataaaatg cctgttactt cctgttactt 3180

```

<210> 370
 <211> 1546
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003862

```

<400> 370
cacggccgga gagacgcgga ggaggagaca tgagccggcg ggcgccagga cggagcggcc 60
gtgacgcttt cgcctgtcag ccgcgcgcc ccagcccccga cgctgtgacc ctggcccacc 120
gcagctccgc cgcggggccg gagagcgcaa cctgcgttcc agaccgccgc gcagctgtgt 180
ccccgactgc agccgggcag ccagcctccc acggagccgc ggacggccgc ccggcagca 240
gtgagctgac tccccgcac cggccaggcg cctcctgcac agcggctcac ccccgcagc 300
ccctgcgccca gccggagagg cgacagcgtc gggagaggcc gcgcggggcg ctgatgcccg 360
agggccgcgc gcggagcgcc ccggagcagc agagtctgca agagtctgca gaggcgccc 420
gggagcagca gcagcggcgg ccggcgccgg ggcgcggcg gagecgccgc ctccctccc 480
cggagcggcg acatgtgca gctgggctag gacttcgcga tgtttacact tccactgtga 540
tattcagccc cctccgcctg caactgtcct ggagaacgtg gacttcgcga accagctcta 600
gtacagctg ttgttgccga cctgaagcag ctgcggctgt accagctcta gacccggacc 660
cgggctcggg acgatgtgag cctgggcccgc aggtatcagt cccgcggcga ggtatgggac 720
agtgggaaac acatccaggt ccgttaagcag ccttcctgga gtcacagctc gatccggccg 780
aagtatgcc acctcctagt ggagacagac gtgcataaac aaggttctg agaacaacta caagccgat 840
aaggagcgg aattctacct gttcatcgag cttgtacgtg ggttccacca agaagggcg ccgcggaa 900
ggcaccagca aggaggtgtg atgtactcgg ccagcaggac ctaagcttca gtaagcgcta ccccaaggg 960
gtgctcgcta agtactcgg ccgggagaa cttcaagacc cgtcggatc cctggcacac 1020
ggcccacaaga cccggagaa cccctgcta ggcaccccgc ccgcggccc tgaagcgtca ccccaaggg 1080
cagcgggagc ttcagaagcc cctcgccta atcagagaa tatttttaca tgaataatga cctggcacac 1140
cggcccacac accctgcta ggcaccccgc ccgcggccc tgaataatga ggaagaact 1200
tcacactccc agaaaactgc atcagagaa ttttttaca tgaataatga ggaagaact 1260
ctatttttgt acattgtgt taaaagaaga caaaaactga accaaaactc ttggggggag 1320

```

```

gggtgataag gattttattg ttgacttgaa acccccgatg acaaaagact cacgcaaagg 1380
gactgtagtc aaccocacagg tgcttgtctc tctctagtaa cagacaactc taaactcgtc 1440
ccagaggagg gacttgaatg aggaaaccaa cactttgaga aaccaaagtc ctttttccca 1500
aagggtctctga aaggaaaaaa aaaaaaaaaa aaaaaa 1546

```

<210> 371
 <211> 376
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig32125_RC

```

<400> 371
caatttgtat tttttcfaat ataatgttgc aagatgtcat tatattctca ttaccataa 60
cttcccatcc aaatcttfaat ggcacatttg gtattttttc aaatggcttt gagatataat 120
tcacatacca cacaattcat gtatctaaag tatcacagttc tgtagtctctg tgggttgttgg 180
ttacatccac cagaactgagg gctccctgag gcttgggtgc tcagctttcc cttcacttac 240
tcattcacca aacattcagc gctactgag acatgggttg atgatggcca acaggagaga 300
cagttctctg ctactgttgg gtgggggtgg agtacttagt aaccgtctgc attatttgat 360
gcttacctgt gctggg 376

```

<210> 372
 <211> 1639
 <212> DNA
 <213> Homo sapiens

<300>
 <308> U82987

```

<400> 372
ccgccccctc cgccgtgttc atgcccccg ggccccaggg agcgccatgg cccgcgcacg 60
ccaggaggcg agctccccgg agcccgtaga gggccctggc cgcgacggcc cgcgccctt 120
cccgctcgcc cgccgtgttc cctcgccagt gtctcgccgc ctctcgcgag ccggcctggc 180
tgccgccccg gccgccccca cctcgctgcc cgctgcctac ctctcgccgc ccacgcgcc 240
accgcccgtc accgcccggc cctgggggttc cgctggcctt ggggggtccc gcagccggcc 300
ccgagggccg cgcccgagcg gtctctcagc ctcgctctcg ctggcgggag agcacctgga 360
gtcgcccggt cccagcgccc cgggggctct ggccggcggt cccaccaagg cggcccggg 420
agtcgcccgg gaggaggaac agtgggccc ggagatcggg gccacagtc ggccgatggc 480
ggagcacttc aacgcacagt acgagcgccg gagacaagag gagcagcag gccacggcc 540
ctcacccctg agggctcctgt acaattctcat catgggactc ctgcccctac ccaggggcca 600
cagagccccc gagatggagc ccaattaggt gcctgcaccc ccccggtgga cgtcagggac 660
tcggggggga ggccccctcc accctctgac accctggcca gcgcggggga ctttctctgc 720
accatgttag atactggaat cccagccctg cctgtcccgg gggcgggcgg ggccagccac 780
tcacgcccga gccacgacct ggggtgcaact actgagatgc ggactcctgg tctccctggc 840
aagaagccag gagaggagcg gctgatggac tcagcatcgg aaggtggcgg tgaccaggg 900
gggtggggact gagccggccc cctctgcccgc ccaccacat ctcaggaaaag gctgtgtgag 960
tgggtcccgt tcagctgca ggggtgacac tggggggggg gggggctctc ctctcggtgc 1020
tccttcaact tggggctggc ctacagcccc tgggtccttc cccctcctc ctgggagggg 1080
gcocgtgaag agcaaatgag ccaaacgtga ccaactagct cctggagcca gagagtggg 1140
ctcgttctct ggttctccca gcccgccgcc cagccatctt cctgagcca gctttagag 1200
gggtggcgat cctgctctac cttoatcagc ggggtggcag gagggggcca gactgtgaat 1260
cctgtgctct gcccgtagcc gcccccgcgc ccatcaatcc cattgcatg gtttagagag 1320
agcagctgtg accactggca ttcatttggg ggggtgggaga ttttggctga agccgcccc 1380
gccttagtcc ccagggccaa gcgctggggg gaagacgggg agtcagggag ggggggaaat 1440
ctcggaagag ggagagtgct gggagtgagg agggatggcc cagcctgtaa gatactgtat 1500
atgctgctgt tagataccg gaatgaattt tctgtacatg ttttggtaat ttttttcta 1560
catgattttt gtatgtttcc ttttcaataa aatcagattg gaacagtgga aaaaaaggcc 1620
ttcgtggcct ccagagatgc 1639

```

<210> 373

<211> 2273

<212> DNA

<213> Homo sapiens

<300>

<308> AL137718

<400> 373

```

ggggggccagt  taccatattg  tccggagtg  gtccttgaca  ttcaggaaag  atacatcctc  60
agacctttcca  aaatttgatc  tgcctctcta  acctaaatac  gccctcatta  aatgcaacttt  120
aatccgagca  ctgtatggct  tggaaagaat  tatgagtgag  gagaggagcc  ttcccttatt  180
ggccaaagcc  gtgatccca  gacaccccaa  tatgatgaca  gatgtggtta  aactctctcc  240
tgcggtatgc  attgtagggg  aagaagacat  ccttgaagaa  gttttagaag  ctttaactct  300
agctcgatgaa  aaaaaaataa  ttgacagatt  tttttgtatt  gtggaaagcc  cctggcaca  360
ttcagttcaa  ctgcaagtag  cttgtatgca  gctcatcaat  gccctgggtt  catctcctga  420
tgatttggat  ttcaggcttc  acatcagaaa  tgaatttatg  cggttggtat  tgaagagagt  480
attgccaaat  ttaaaatgca  ttaagaatga  tggcctggat  atcccaacta  aagtcctttga  540
tgagcataaa  gaagaagatt  tgtttgagtt  atcccatcgc  ctgtaagata  ttgagactga  600
acttgatgaa  gcatatgatg  ttacaacat  ggtgtggagc  acagttaaag  aaactagagc  660
agaggggat  tttatttcta  ttcttcagca  tcttttgctg  attcgaattg  attattttat  720
aaggcaacaa  tacttcaaat  taattgatga  gtgtgtatcc  cagattgtat  tgcatagaga  780
tggaattggat  ccagacttca  catatcgaaa  aagactagat  ttgatttaa  cccagtttgt  840
agacatttgc  atagatcaag  caaaactaga  agagtttgaa  gagaagcatt  cagaacttta  900
caagaaattt  gaaaaagagt  ttaccgacca  ccaagaacct  caggctgaat  tgcagaaaaa  960
agagcgcaag  attaatgagc  ttcaagcaga  gctacaagct  ttttaagtct  agtttggctg  1020
cttgccagct  gattgttaata  ttcttttgc  tctctctaaa  gaaggtggaa  ctggccactc  1080
agcacttctc  cctccgctcc  cactgccttc  gtgctggagg  gtgcgcctcc  cactcctccc  1140
cccaccact  cctccacttc  caggaatcgc  gatgccattc  agtggctcct  tgctccacc  1200
acctccctgc  ggattccttg  gaggacaaaa  ttctctctcc  ctaccaatcc  tggcatttgg  1260
gttgaaacca  aagaaagaat  ttaaacctga  aatcagcatg  agaagattga  atttggttaa  1320
gatcagacgt  catgaaatga  ctgaaaaactg  tttctggata  aagtaaatg  aaaaatagta  1380
tgaaaacgtg  gatttgcttt  gtaacctga  gaatcacatt  tgtgccaac  aaaaagagag  1440
aagagaagag  gaagatattg  aagaaagaa  atcgattaa  aaaaaatta  aagaacttaa  1500
gtttttagat  tctaaaaattg  ccagaaacct  ttcaatcttc  ctgagctctt  ttcgggtgcc  1560
atatggaggaa  atcagaatga  tgatatttga  agtagatgaa  acacggttgg  cagagtctat  1620
gattcagaac  ttaataaagc  atcttctga  tcaagagcaa  tttaattcat  tgcctcagtt  1680
caagagtgaa  tatagcaact  tatgtgaacc  tgagcagttt  gtggttgtga  tgagcaattg  1740
gaagagacta  cggccacggc  tcaagtctat  tctctttaa  cttcagttt  aagagcaggt  1800
gaacaacatc  aaacctgaca  tcatggctgt  cagtactgcc  tgcgaagaga  taaagaagag  1860
caaaagcttt  agcaagttgc  tggaaactgt  attgtcaatg  ggaactacta  taagtgctgt  1920
ctcccggaat  gctcaaacct  tcggatttaa  ccttagctct  ctctgtaaac  taaaggacac  1980
aaaaacgaca  gatcagaata  caacgctact  tcaattctgc  gtagaatat  gtgaagagaa  2040
ttacctgat  atactgaatt  ttgtggatga  tttggaacct  tttagacaa  ctgataaagg  2100
gttgctcctt  tttagaatac  atttcagggc  tctgactctc  agtgctaaaa  gactgaaaaa  2160
tataaccttt  atttgtatgt  actttccact  ttctcacagt  gttttcatac  ccaatatctc  2220
attttaaaaa  taaagagagt  tgtaggatat  taaaaaaaaa  aaaaaaaaaa  aag  2273

```

<210> 374

<211> 2782

<212> DNA

<213> Homo sapiens

<300>

<308> AB037863

<400> 374

```

cgggccggcg  cgggcgggcg  cggaacccg  gcccggggg  gactggcgcg  ggctgtgtct  60
gctgtgtctc  caggtgctgc  cgcccgggcg  tacggcagcg  ccgggacgcg  ggggcaacgc  120
gggcgtggcg  ggccgctggc  tctgtgtggc  cgcccgggcg  ccagccttct  ccgggagcgc  180
ggcgctgtgc  ctgctcctgc  cgccgctgcc  gccactgtcg  ccgcgcgcgc  cgagctccgc  240
gcccgacgc  tccgcctccc  ggatggacgc  tctgcccgc  agcgggtga  acctgaagga  300

```

```

ggagcgctg ctgcccgcg gcctgggctc agtgcgctc tggatgcagg gcgggggcat 360
ctgggacgac agcacgcgag cgcagagtg gctgggtctg gcacagagac attttgagaa 420
gcagcctccc tccaacctca ggaataccaa cttcttccac ttcgtgctgg ccatgtacga 480
ccggcagggg cagccctgag aggtggagcg cacagccttc atcgacttcg tggaaaagga 540
ccgagagccc gggcgggaaa agactaacaa tgggattccat tacgcctccg ggcctgggta 600
taacaatgga ctgcggacag agcaagacct ctacgtgcgt ctcacgcact ccatgtccaa 660
acaggccatc atctatgagg ggacagacaa gaaccccgaa atgtgcgag lgcgtgtcac 720
ccatgagatc atgtgcagcc ggtgctgtga ccggaaagag tgtggcaacc tggatggagac 780
gcctccagac cccgtcatca ttgacaggtt cttcctcaag ttcttccata aatgcacaa 840
gaactgcctg aagaatgcgg ggaatccag agacatgcgc cgttccagg tgggtggtgc 900
cacagcggtg agcgtggagc gacacgtgct ggcggtgtcc gacaacatgt ttgtgcacaa 960
caactccaag catggccgca gggcgccgag cctggaaccc tccgaagct ccaacccctg 1020
catcaaggcg tctcccccgc gggagggctg gaccacgggc ggcgccaccg tcattgtcat 1080
cgcgacaac atcttgcagc ggttgcaggt cgtgttcgga aacgtgctcg tgtggagcga 1140
gctcatcagc cccacgcgca tccgggtgca gacgccccgc cgccacatcc ccgggggtgt 1200
ggaggtgacc ctctctaca agtccaagca gttttgcaa ggtatgcccc gccgctttgt 1260
ctacacagct ctgaacgagc ccaccattga ctacggattc gccaaggaa gtgctgctga agcggcgccg 1320
tccagacac cccggagacc ccgagaggtc gcccaagtaa caggagctac agaaagtcat 1380
cgacttgga gaagccctgt acggagtgcc cggcagtaac caggagctgc tctctgaagc 1440
ccggcgagag gtggccgagg ctctgtacag cacccccgc gcacccgggc cgtcgcaccc 1500
cgtggcccg agccacccac actccgcgt gttgggcatc aacgcttca gacgccctg 1560
ggccatcgcc ctggggagc ccaccccggg gcccgagccg ggctacgcgc gcagctgca 1620
cagcgctgcc ccccgcggtg tcgcgcccaag ccccgctcgc cagcagagcg gtacggcg 1680
cgccctcgga gctggctcg gcggctacgg cgcgcggggc gtggccggcc tcggctgctc 1740
tgggtcccc agcttccata atggctccac cccttcgca ccttcgca cgaagcgccct 1800
tcgccccgct gctgcgcccc ccaagctccc caccacggc ctgccccaga gccacggag 1860
aggggcttcc agaccagct tttgagatt ctgacaagtt ccaactctca ccggcggggc 1920
ttcaggccct ggcatactcc taattacggt ctgcagctgt tccatggag ccgggactgg 1980
aggtccctct gggattcaca gccacacccc ggaatgtggc acagacagat gaaggccac 2040
ggccatgggg ggaacctcaac ccgtgagctg aacggggaga ggccttacc ccatgtcaa 2100
gcctccccgc tagcagcccc acaggcttct ctgcctccc tgtcttggg tagtcagaa 2160
ccccagcact gtgcagatgc tcttggcagg acagatcgc agggaggtgc tgggattctg 2220
ggcctcactg tctgggtctt ggttccctg aaagagatg atcttgtca gccagggtt 2280
gttgagtga gggagcgtgg gatggggacc tggggaaaag ggcagctca gggagaagt 2340
acctgaaaag gtctgtttg catctgacc atctcaactg gccacgcat ccaacttctc 2400
tgacgcgaaa ggtgtggccc ccgcagcctc gggaggctcc cccaggctcc cgtggagctt 2460
ccacacagct cttggccccg cagctgcccc cacttccctt gagactgca ctctcatgtc 2520
tgccgcatca tcttccctg tgggggctt ggccatggag gaggcagaag aggggggtgc 2580
agggctcctg tatttggggt ctcccccaag tggatgtctc atggactctg gccccaacaa 2640
ctcaaatgca ctctggctgg cccacgcag cggccccag ccgcccccag gtcggctcac 2700
attctgctct gctaagtttg gagaaaacag aacaataaac catgacagg tgggtccccc 2760
ccggcctctc acctgcctcc tt

```

<210> 375
<211> 716
<212> DNA
<213> Homo sapiens

<300>
<308> NM_020188

```

<400> 375
acgggtgagc ggtggagggc gtccactgggt ttcgggtct ggcaagcggt tcagctgtct 60
gtcccttagc agccggcctt cgggtcgggc gtcctcgcgc ctactgcgcg ttcagttctc 120
ccggtgtggc cagcagctgg gtgctactgc tctgatccat cctcatctcc taaagatgca 180
tctctactta ctccacact tgccacactga agaatagcaac gtcttgatta acttgcttaa 240
ggaatgtcac aaaaatcaca acattctgaa atttttgggt tattgtaatg actgtgatcg 300
ggagtgtgaa aaatgcctga agaataagta cgtagaaaac aggaaccaaga gcaggagaga 360
tgcatttgca atgcgaaaga aactttttaa tctccagag gaatccgaaa aataaattgt 420
attttcaact gatgccttg ctgagagaa acctaaagac tctgggttga taactgaaag 480
aatctgtctc tatttgggtc ccaataactc ttgaatggaa agtgacctgt gagagattga 540

```

accatggaga aatatgaaaa cctctgattc tgagtatttg ttgggcaggg cgttttagtac 600
 tgtctccctc ttaccagcaa acctgacttc accatgttta ttccctttgc ctacacacag 660
 ttaatactcg agtaacttat ctcttcaat aaaataattt aaataaaaaa aaaaaa 716

<210> 376
 <211> 3737
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_020974

<400> 376
 ggcgctccgg cacacctccc cgcgccggcg cggccaccgc cggcactccg cggcctctgc 60
 cgcacaaccg tgagccatcc atgggggtcg cgggcccga cggctccggg cgggctccgg 120
 cgggtgctgc gctgctgctg ctgctgcgcg cactgctgct gctggcgggg ccgctcccgc 180
 cgggtcgggg cgtgcccgcg gggccgcagg aggatgtaga tgagtgtgcc caagggtcag 240
 atgactgccca tgcgaagcgc ctgtgtcaga acacacccac ctctacacag tgctctgcga 300
 agcctgggcta ccaaggggaa ggcaggcagt gtgaggacat cgtacacgcg tcggaagagg 360
 tcaatggagg ctgtgtccat gactgtttga atattccagg caattatcgt tgcaatttgt 420
 ttgatggctt catgttggtc catgacggtc ataatgtctc taatgtggac gagtgcctgg 480
 agaacaatgg cggctgccag catacctgtg tcaacgtcat ggggagctat gagtgtcgtc 540
 gcaaggaggg gtttttctg agtgacaatc agcacacctg cattcacgcg tcggaagagg 600
 gctgtgagct catgaataag gatcacggct gtatgcacat ctgcaaggag gccccaaggg 660
 gcaagctcgc ctgtgagtcg aggcctgggt ttgagctggc caagaaccag agagactgca 720
 tcttgacctg taacctaggg aacggtgggt gccagcactc ctgtgacgat acagccgatg 780
 gccagagtg ccagctggat ccacagtaca agatgcacac agatgggagg agctgccttg 840
 agcgagagga gactgtcctg gaggtgacag agagcaacac cacatcagtg gtggatgggg 900
 ataaacgggt gaaacggcgg ctgctcatgg aaacgtgtgc tgccaacat agtggaggtg 960
 cccgcacctg taaggatact tgcacagggt tccactgcag ttgtcctgtt ggattcactc 1020
 tccagttgga tgggaagaca tgtaaagata ttgatgagtg ccagaccgcg aatggaggtt 1080
 gtgatcattt ctgcataaac atctgtgggca gttttgactg cggctgcaag aaaggattta 1140
 aattattaac agatgagaag tcttgccaag atgtggatga gtgctcttgg gatagacct 1200
 gtgaccaacg ctgcatcaac caccctggca catttgcttg tcttgccaac cgagggtaca 1260
 cctctgatgg ctccacccac tgtggagaca ccaatgagtg cagcatcaac aacggaggtc 1320
 gtccagaggt ctgtgtgaac acagtgggca gctatgaatg ccagtggcag cctgggtaca 1380
 agctccactg gaataaaaaa gactgtgtgg aagtgaaggg gctcctgccc acaagtgtgt 1440
 caccctgtgt gtccctgcac tgcggttaaga gtggtggagg agacgggtgc tctctcagat 1500
 gtcaactctg cattcacctc tcttcagatg taccaccact caggacaagt gtaaccttta 1560
 agctaataag agccaagtgt agtttgaaaa atgctgagct gtctcccgag gtctgcgac 1620
 cagcactacc agagaagcac agctcagtaa aagagagctt ccgtacgta aaccttaac 1680
 cagctctgct caagcagatc ccaggagccc ctggccgacc aagcacctc taggaaatgt 1740
 ttatcactgt tgaactgttg cttgaaacta accaaaaagg ggtgacagct tcttgtagcc 1800
 tgagctgcat cgtaaagcga accgagaagc agctcogtaa agcatccgc agctcagaa 1860
 agggcgtcca cagggacag tttcacctcc agctctcagg catgaacctc gacgtgacta 1920
 aaaaagctcc cagaacatct gaacgccaag cagagtccgt tggagtgagg cagggtcatg 1980
 cagaaaaacca atgtgtcagt tgcagggctg ggaacctatta tgatggagca cgagaacgct 2040
 gccattttat tccaaatgga accttccaaa atgaggaagg tgtgaacctg tgtgaacct 2100
 gcccaagacc agaaaattct ggggccctga agaccccaga agcttggaa atgtctgaat 2160
 gtggaggctct gtgtcaacct ggtgaattt ctgcagatgg ctttgcaact tgcagctct 2220
 gtgacctggg caegtccag cctgaagctg ctggaacttc ctttccccc gtgagggag 2280
 gccctggcac caaacatcag ggagctactt cctttcagga ctgtgaaacc agagtcaat 2340
 gttcacctcg accattctac aacaccca ctcaaccgat tattctgtgc ccagtgaggaa 2400
 catcacagcc tgaatttgga aaaaaataat gtgtttcttg ccaaggaaat actacgactg 2460
 acttttagtg ctccacaac ataacccagt gtaaaaaag aagatgtgca ggggagctgg 2520
 gagatttcaac tgggtacatt gaatcccaa actaccagg caattaccca gcccaaccg 2580
 agtgtacgtg gacctcaac ccaaccccca agcgcgcgat cctgatctgt cctctgaga 2640
 tcttctcgtg catagaggac gactgtgggg actatctggt gatcgggaaa acctcttact 2700
 ccaattctgc gacaacatat gaaacctgoc agacctacga agcggcctc gctgtacct 2760
 ccaggtcaaa gaagctgtgg attcagttca agtccaatga agggaaacag cgtgaggggt 2820
 tccaggtccc atacgtgaca tatgatgagg actaccagga actcattgaa gacatagttc 2880

gagatggcag	gctctatgca	tctgagaacc	atcaggaaat	acttaaggat	aagaaactta	2940
tcaaggctct	gtttgatgtc	ctggcccatc	ccagagaact	tttcaagtac	acagcccagg	3000
agtcccgaga	gatgtttcca	agatcgttca	tccgattgct	acgttccaaa	gtgtccaggc	3060
ttttgagacc	tacacaaatga	ctcagcccat	gtgccactca	atacaaatgt	ctctctatag	3120
ggttggtggg	acagagctgt	cttcctctct	catgtcactga	cagtcgggta	ttgtcgcctc	3180
cgtatcatag	cattcattag	agttcaattt	ttatagataa	tacagatatt	ttgttaaat	3240
gaacttggtt	tttctttccc	agcatcgtgg	atgtagactg	agaatggctt	tgagtggcat	3300
cagctctctca	ctgctgtggg	cggatgtctt	gcgatagatca	cgggctggct	gagctggact	3360
ctgtgcagcc	taggtgagac	tcacctgtcc	ttctggggct	ttactcctcc	tcaaggagtc	3420
tgtagtggaa	aggaggccac	agaataagct	gcttattctc	aaacttcacg	ctctctatag	3480
ccggccctct	ctaaggggag	cctctgcact	cgtgtgcagg	ctctgaccag	gcagaaacag	3540
caagagggga	gggaaggaga	ccctgcagg	ctccctccac	ccaccttgag	actctggagg	3600
actcagtttc	tccacagcct	tctccagcct	gtgtgatata	agtttgatcc	caggaaactg	3660
agttctaagc	agtgctcgtg	aaaaaaaaa	gcagaaagaa	ttagaaataa	ataaaaaacta	3720
agcaettctg	gagacat					3737

<210> 377
 <211> 3304
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000127

<400> 377						
ggcagccgaac	ggggcggtcg	gcagcgttgc	cgcgggggcc	tgcgaagcgc	tgctcggggc	60
cgccactcgc	cgcggggagg	acgcgcgcgc	gcgcgcaccc	agcgccgcgc	cgccgcgcgc	120
ctccagccgg	gcgcgcgcgc	gtcccggggg	ccggcccccgc	gagcgccagg	gtaaacacgc	180
ccggagtcctt	ggagcgcgctg	cagaaggaaa	taaaagagaga	tgacgggatt	ttbgaggtta	240
cgggcgcccca	gctgcgaagat	gcactagccg	gctgaacccg	ggatcggtgc	actgtgttga	300
accggagtgct	ctctgcacgga	gagtggttga	tgagttgaag	ttgccttccc	ggggctcatt	360
ttccagtgtaa	ccgagaggaa	tcgagagggc	aaggccaatca	ctctgctctg	ccattgattg	420
ggfatcgagg	gctttttttt	tctccctctc	ctcttttttt	tcctccgtct	tggtgcattgc	480
aagaaaaatt	gactccgctg	ctcgcccgcc	ctgggtgcga	gatattcagc	ccgcctctct	540
cccgctgcat	gtgcaaccca	aagatgaaag	accgaagggg	agaaagttaa	agaaatcgcc	600
cacatgcgct	ggatcagtc	acggcttggg	gaaaggcatc	cagagaaggc	gggagcggag	660
agtttggaagt	ctttacagcg	gggaagatgg	cggactggag	ctgaaagtgt	tgattgggaa	720
acttggttga	ttctgtgttt	tatttacaat	ctctttgacc	caggcaggac	acatcgagcg	780
caaaaaacgc	tatttcaatc	tgctctcagc	tggtctctgt	ctcgcccttt	tggtttattt	840
cgaggcgcttg	cagtttaggg	catcgaggag	ccacagccgg	agagaagaac	acagcggtag	900
gaatgctgtg	caccacccca	gtccggtatc	tttctggccc	cgcttccggc	agcctctgcg	960
ccctcttgctt	ctctggggtc	aattggaaaa	cgaggattcc	agcgtgcaca	tttccccccc	1020
gcagaaacga	gactgccaat	ccagcatcta	caaaggcaag	aagtggccga	tgagggtctg	1080
cttcgatttc	accctttgca	agaaaaacgg	cttcaaagtc	tacgtatacc	ccagcaaaaa	1140
aggggagaaa	atccgcgaaa	gttaccacaa	cattctagcg	gccatcgagg	gctccaggtt	1200
ctacacccct	gcacccagcc	aggggtgcct	ctttgtcctg	agctctgaga	ctttagacag	1260
agacagcttg	tcaoctcagt	atgtgcacaa	tttgagatcc	aaagtcgaga	gtctccactt	1320
gtggaaacat	ggtaggaatc	atttaatttt	taatttatat	tcggcgactt	ggcctgacta	1380
caccgaggac	tggggggttg	acatcgccca	ggcgatgctg	gccaaagcga	gcactcagta	1440
tgaaaaactc	cgaaccaact	ttgatgtttc	tattccccct	ttttctaaag	atcatccagg	1500
gacagagggg	gagagggggg	ttttgaagtt	caacaccatc	ctctctctca	ggcagtcact	1560
gctggtattc	aagggggaaga	ggtacctgac	agggtagtag	ctcagaccca	ggaatgcctt	1620
atacagctc	ctaacggggg	aggacgttgt	ctctctcacc	acctgcaagc	atggcaaaag	1680
ctggcaaaag	cacaaggatt	ctcgtctgtg	cagagacaac	accgagtatg	agaaagtatg	1740
ttatcgggaa	atgctgcaca	atgcccacttt	ctgctctggtt	ctctgctggtc	cgaggctgtg	1800
gtcctctcaga	ttctcggagg	ctttgcaggc	tgctctcgctc	ctcgtgtatgc	tcaggcaatg	1860
atggaggttg	ctactctctc	aagtgattaa	ttggaaacca	gctgcgctga	cagcggtatg	1920
gagatgttta	ttacagattc	ctttacaaat	caggtctatt	catcaggata	aaatcctagc	1980
acttagacag	cagacacaat	tcttgtggga	ggcttatttt	tcttcagttg	agaagattgt	2040
attaactaca	ctagagatta	ttcaggacag	aatattcaag	cacatatcac	gtaacagttt	2100
aatatgggac	aaacatcctg	gaggatttgt	cgtactacca	cagttattcat	cttactctgg	2160

```

agattttcct tactactatg ctaatttagg tttaaagccc ccctccaaat ttoactgcagt 2220
catccatgcg gtgacccccc tggctctcta gtcccagcca gtgttgaagc ttctcgtggc 2280
tcgagccaag tcccagctact gtgccagatg catagttcta tggaaattgtg acaagccctc 2340
accagccaaa caccgctggc tgcgcactgc tgtgcctgtc tgcgtcattg aaggagagag 2400
caaggttatg agcagccgtt ttctgcccta cgacaacatc atcacagacg cgtgctgcag 2460
ccttgacgag gacacggtgc ttcaacaac agaggtggat ttgccttcoa cagtgtggca 2520
gagctccctc gagaggattg tggggtaccc cgcgcgcagc cacttctggg ataactctaa 2580
ggagcgtggg gatacacat caaagtggac gaacgactac tccatggtgt tgacaggagc 2640
tgctatttac cacaaatatt atcactacct atactcccat taccctgccag ccagcctgaa 2700
gaacatggtg gaccaattgg ccaattgtga ggacattctc atgaacttcc ttgtgtctgc 2760
tgtgacaaaa ttgcctccaa tcaagtgcac ccagaagaag cagtataaag agacaatgat 2820
gggacagact tctcgggctt cccgttgggc tgacctgcac cactttgcc agcgacagag 2880
ctgcatgaat acgtttgcc a gctggtttgg ctacatgccg ctgatccact ctgcatgag 2940
gtcgcacccc gtctcttcta aagaccaggt ctctattttg aggaagaat accgagacat 3000
tgagcacttt tgaggatacc ggctgagttg gggaggggaa gcaagaaggg atggggtgca 3060
agctcgtctc tcttcccagt gcagatccac tcatcagcag agccagattg tgccaactct 3120
ccaaaacttt agatgcagc aatgacaaaa aaaaaaaagg ccaatgagaa ctaactctct 3180
ggctcctggg actgcacag actgctccaa actcacctca ctggcttctg tgtcccaaga 3240
ctaggttggt acagtttaat tatggaacat taataatata ttttgaaaaa aaaaaaaa 3300
aaaa 3304

```

<210> 378

<211> 7680

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002019

<400> 378

```

gcggacactc ctctcggctc ctccccggca gggcgggcgg ctgggagcgg gctcgggggc 60
ctcgggtcag cggccagcgg gcttggcgcg gaggattacc cggggaagtg gttgtctcct 120
ggctggagcc gcgagacggg cgctcagggc gcggggcgcg cgcgcgcgaa cgagagagac 180
gactctggcg gccgggtcgt tggcgggggg agcgcgggca cggcgcgagc agggcggctc 240
gcgctcacca tggctcagcta ctgggacacc ggggtcctgc tgtgcgcgct gctcagctgt 300
ctgcttctca caggatctag ttcaaggttca aaattaaaag atcctgaact gaggttaaaa 360
ggcaccaccg acatcatgca agcaggccag acactgcatc tccaatgcag gggggaagca 420
gccataaatt ggtctttgcc tgaatgggtg agtaaggaaa gcgaaggcct gacataact 480
aaatctgcct gtgggaagaa tggcaaacaa ttctgcagta ctttaacctt gaacacagct 540
caagcaaacac acactggctt ctacagctcg aaatatctag ctgtacctac ctgtacactc 600
aaggaaacag aatctgcaat ctatatattt attagtata caggtagacc tttcgtagag 660
atgtacagtg aaatccccga aattatacac atgaactga atgaactgag gaaggagct 720
tgccgggtta cgtcacctaa catcactgtt atgatctat ctgtacctac ctgtacactc 780
atccctgatg gaaacgcgat aatctgggac agtgaaaagg gcttcatcat atcaaattgca 840
acgtacaaag aaataggctt tctgacctgt caataggcca tcaattggca ttgtataag 900
acaaactatc tcacacatcg acaaaccaat acaatcatag atgtccaaat aagcacacca 960
cgcccagtc aattactttag aggcctactc atgttactgc ctgtctctca tccactccc 1020
ttgaacacga gatttcaaat gacctggagt taccctgatg aaaaaataa gagagcttcc 1080
gtaaggcgac gaattgacaa aagcaattcc cacttcaaca tattctacag tttcttact 1140
attgacaaaa tgcagacaaa agacaaaagg aattatactt gtctgttaag gagtggacca 1200
tcattcaaat cgtttaaacc ctcaagtcat atatatgata aagcattcat gttctgaaa 1260
catcgaaaac agcagctgtc tgaaaaccta gctggcaagc ggtcttaccg gctctctatg 1320
aaagtgaagg catcttccctc gccggaagtt gctatggtta aagatgggtt acctgcgact 1380
gagaaatctc ctgcctattt gactcgtggc tactcgttaa ttatcaagga cgttaactga 1440
gaggatcgag ggaattatcc aatcttctg atcataaaac agtcaaatgt gtttaaaact 1500
ctcactgcca ctctaattg caatgtgaaa cccagattt acgaaaaggc cgtgtcatcg 1560
tttccagacc cggctctcta cccactgggc agcagacaaa tctcgacttg taccgcata 1620
ggtatccctc aacctacaat caagtgttc tggcaccctc gtaaccataa tcatccgaa 1680
gcaaggtgtg actttgttgc caataattga gactccttta tccctgatgc tgacagcaac 1740
atgggaaaca gaattgagag catcactcag cgcattggca taattagaag aaagataaag 1800
atggctagca ccttggttgt ggctgactct agaatttctg gaattctacat ttgcatagtc 1860

```

tccaataaag	ttgggactgt	gggaagaaac	ataagctttt	atatcacaga	tgtgccaaat	1920
gggttttcag	ttaacttgga	aaaaatgccg	acggaaggag	aggacctgaa	actgtctctgc	1980
acagttaaaca	agttcttata	cagagacggt	acttggattt	tactggcgag	agtttaataac	2040
agaacaatgc	actacagtat	tagcaagcaa	aaaatggcca	tcactaagga	gcactccatc	2100
actcttaatc	ttacatcat	gaatgtttcc	ctgcaagatt	cagggcacct	tgcctgcaga	2160
gccgaagaag	tattacaagg	ggagaagaatc	ctccagaaga	aagaaattac	aatcagagat	2220
caggaagcac	ctactcctct	gcgaaacctc	agtgatcaca	cagtggccat	cagcagttcc	2280
accactttag	actgtctatgc	taatgggtgc	cccgagcctc	agatcacttg	gtttaaaaac	2340
aaccacaaaa	tacaacaaga	gcctggaatt	attttaggac	caggaagcag	cagcgtgttt	2400
attgaagag	tcacaagaag	ggatgaaggt	gtctatcaat	gcaaacccac	caaccagaag	2460
ggctctgtgg	aaagttcagc	atacctcact	gttcaaggaa	cctcggacaa	gtctaatctg	2520
gagctgatca	ctctaacatg	cacctgtgtg	gctgcgactc	ctctctggct	ctattaaacc	2580
ctccttatcc	gaaaaatgaa	aaggtctctc	tctgaaataa	agactgacta	ccatcaatt	2640
ataatggacc	cagatgaagt	tcctttggaat	gagcagtggt	agcggtctccc	ttatgatgcc	2700
agcaagtggg	agtttggccg	ggagagactt	aaactgggca	aatcacttgg	aagaggggct	2760
tttggaagaag	tggtttcaagc	atcagcattt	ggcatataa	aatcacctac	gtgcggagct	2820
gtggctgtga	aaaatgctgaa	agagggggcc	acggccagcg	agtcacaagc	ctctgatgact	2880
gagctcaaaa	tcttgaccca	cattggccac	catctgaacg	tggttaacct	gctgggagct	2940
tgcccaaacg	aaaggaggcc	ctctgatgtg	attgttgaat	actgcaaat	tgtaaatctc	3000
tcaccaactc	tcaagagcaa	acgtgactta	tttttctc	acaaggatgc	agcactacac	3060
atggagccta	agaaagaaaa	aatggagcga	ggcctggaac	aggccaagaa	accagaacta	3120
gatagcgtga	ccagcagcga	aagctttgog	agctccggct	ttcaggaaga	taaaagtctg	3180
agtgaatttg	aggaagagga	ggattctgac	ggtttctaca	aggagcccat	cactatggaa	3240
gatctgattt	cttaccagtt	tcaagtgccc	agaggcatgg	agttctctgc	tccagaagag	3300
tgcatctact	gggacctggc	agcgagaaac	attcttttat	ctgagaacaa	cgtgtgtgaag	3360
attttgtatt	ttggccttgc	cogggatatt	tataagaacc	ccgatattgt	gagaaaaagg	3420
gatactcgac	ttcctctgaa	atggatggct	ccogaatcta	tctttgacaa	actctacagc	3480
accaagagcg	acgtgtggtc	ttacggagta	ttgctgtggg	aaatcttctc	cttaggtggg	3540
tctcctaac	caggagtaca	aatggatgag	gacttttgca	gtcgcctgag	ggaaggcatc	3600
aggatgagag	ctcctgagta	ctctactcct	gaaatctatc	agatcatctg	ggactgtctg	3660
cacagagacc	caaaagaaag	gccaaagatt	gcagaactct	tgaaaaaact	aggtgatttg	3720
cttcaagcaa	attgtacaaca	ggatggtaaa	gactacatcc	caatcaatgc	ctactctgaca	3780
ggaatagtgt	gggttacata	ctcaactcct	gccttctctg	aggacttctt	caaaggaaagt	3840
atttcagctc	cgaagttaa	ttcaggaagc	ctgatgatg	ctagatgtgt	aaatgcttct	3900
aagttcatga	gcctggaaag	aatcaaaacc	tttgaagaac	ttttacogaa	tgccaactcc	3960
atgtttgatg	actacacagg	cgacagcagc	actctgttgg	cctctcccat	gctgaagcgc	4020
ttcacctgga	ctgacagcaa	accacaaggc	tcgtctcaag	ttgacttgag	agtaaccagt	4080
aaaagttaag	agtccgggct	gtctgatgtc	agcagggcca	gtttctgcca	ttccagctgt	4140
gggacgtcta	gcgaagagcaa	gcgcaggttc	acctacgacc	acgctgagct	ggaaaggaaa	4200
atcgctgtgc	gctccccgcc	cccgagctac	aactcgttgg	tctgtacttc	acccccaccc	4260
atctagagtt	tgacacgaag	ccttatttct	agaagcacat	gtgtattttt	acccccagga	4320
aactagcttt	tgccagttat	atgcataat	aagtttacc	ctttatcttt	ccatgggagc	4380
cagctgtctt	tttgtatttt	tttaatagtg	ctttattttt	ttgaactaca	agaaatgtac	4440
tccagtagaga	caaatatgtg	caagtgaaga	acactactgc	taaatctcta	tgttactcag	4500
ttgttagaaa	atccttctcta	aacccaatga	cttccctgct	ccaacccccc	ccacctcagg	4560
gcacgcagga	ccagtttgat	tgaggagctg	cactgatcac	ccaatgcact	acgttaccoca	4620
ctgggcccgc	cctgcagccc	aaaaccaggg	gcgaacagcg	cgtagcccc	aggggactac	4680
tggtctggct	gagcaacatc	tcgggagctc	ctcagcaggc	ctaagacatg	tgaggagaaa	4740
aaggaaaaaa	agcaaaaaagc	aaggggagaaa	aggaagaacc	ggagaagcga	tgagaaagaa	4800
tttgagagc	accatgtggg	cacggagggg	gacggggctc	agcaatgcc	tttcagttgc	4860
ttccagctc	tgacccttct	acatttgagg	gccagccgag	cagcagatg	acagcgatga	4920
ggggacattt	ctctggaatt	gggaggcaag	aaaaggacaa	atatcttttt	tggaactaaa	4980
gcaaatttta	gacctttacc	tatggaaagt	gttctatgct	gttctctatc	cgtagctagt	5040
tttgtattgt	agcaactgag	gtggcactca	actctgagcc	catacttttg	gtcctctag	5100
taagatgcac	tgaaaaactta	gccagagtta	ggttgtctcc	agggccatgt	gctcttacc	5160
tgaagtattc	acattctatt	ttgggtatta	atataatgct	cagacactta	actcaatttc	5220
ttgtatttat	ctctgttttgc	acagtttagt	gtgaaagaaa	ctgtgagaag	atgaaaatgc	5280
agtcctgagg	agagttttct	ccatatcaaa	acgagggctg	atggagaaaa	aaggtcaata	5340
aggtcaaggg	agaaacccgt	ctctatacca	accaaaacca	ttcaccaaca	cagttgggac	5400
ccaaaacaca	ggaagtccgt	cacgtttcct	tttcatttaa	tggggattcc	actatctcac	5460
actaatctga	aaggatgtgg	aagagcatta	gctgggcat	attaagcact	tcaagctcct	5520

tgagtaaaaa	ggtggatgt	aatttatgca	aggatattct	ccagttggga	ctcaggatat	5580
tagttaaata	gcoactcaata	gaagaaaagc	ccatttttcaa	ctgcttttgaa	acttgccctgg	5640
ggtctgagca	tgatgggaat	agggagacag	ggtaggaagc	ggcgccatct	cttcagggttc	5700
taaagatcaa	gtgggctctg	gatcgctaa	ctggctctgt	tgtatgctat	tatatcaagt	5760
tagggctcat	gtatttagga	tgcgcctact	cttcagggtc	taaagatcaa	gtgggctctg	5820
gatcgctcaag	ctggctctgt	ttgatgctat	ttatgcaagt	tagggctcat	gtatttagga	5880
gtgtcgacc	ttctgcagcc	agtcagaagc	tggagagcca	acagtggatt	gctgcttctt	5940
ggggagaaga	gtatgcttcc	ttttatccat	gtaatttaac	tgtagaacct	gagctctaa	6000
taaccgaaga	atgtatgctc	ctgttcttat	gtgccacatc	cttgtttaaa	ggtctctctg	6060
atgaagagat	gggacgcctca	tcagcacatt	cctagtggag	cctactggct	cctgggacgc	6120
gcttttggg	aagactcact	agccagaaga	gaggagtggg	acagtctctc	ccaccaagat	6180
ctaaatccaa	acaaaagcag	gctagagcca	gaagagagga	caaatctttg	ttgttctctc	6240
tottttacaca	tacgcaaac	acctgtgaca	gctggcaatt	ttataaatca	ggtaacttga	6300
agggaggttaa	actcagaaaa	aagaagacct	cagtcaattc	tctacttttt	tttttttttt	6360
tcocaaatagc	ataatagccc	agcaaatagt	gatacaaat	aaaaccttag	ctgttcatgt	6420
cttgatttca	ataattaatt	cttaatcatt	aagagaccat	aataaaactc	ccttttcaag	6480
agaaaagcaa	aacctattaga	attgttactc	agctccttca	aactcaggtt	tgtagcatac	6540
atgagtcct	ccatcagctca	aagaatggtt	ccatctggag	tcttaagtga	gaaagaaaaa	6600
tggagactgt	taataaagag	ctagttacaa	agtgtgtgtt	cattaaaaata	gcaactgaaa	6660
ttgaaacatg	aatttaactga	taataatcca	atcattttgcc	atttatgaca	aaaattgttg	6720
gcactaacaa	agaacgagca	cttcctttca	gagtttctga	gataatgtac	gtggaacagt	6780
ctgggtggaa	tggggctgaa	accatgtgca	agtcctgtgc	tgtcagctgc	aagaagatgc	6840
accgagatgt	taatttttagg	gaccogtgcc	ttgtttccta	ggccacaaga	atgcaaaatg	6900
caaacagata	ctcgctagcc	tcatttaaat	tgtattaaag	aggagtgcac	ctttggccga	6960
cactggtgtga	actgtgtgtg	tgtgtgtgtg	tgtgtgtgtg	gaatttttaa	gttactttta	7020
ggtgtatgtg	tggtttgtgc	ataactattt	aaggaaacctg	ggtttgtgct	tatatcttca	7080
tacaaaccaa	gaatatatgc	tacagatata	agacagacat	tacttaaaaa	tatttctctg	7140
gtcatgatga	atgtattttg	tataccatct	tcatataata	caactgcttt	tttttcttaa	7200
ttgggatttg	taactgtacc	aacttaattg	ataaacttgg	taagctgctt	tattgtctgt	7260
ctctcttcat	aaatttttca	aaatacta	tcaacaaaga	aaaagctctt	tttttcttaa	7320
aaataaaact	aaattttatcc	ttgttttagag	cagagaaaaa	ttacagaaaa	cttttagaatt	7380
gtctcaaaaa	attgtctaat	attttcaatg	gaaaactaaa	tgttagttta	gctgtattga	7440
tgggggtttc	gaacctttca	ctttttgttt	gttttaacct	tttcaacctc	gtgtaattg	7500
coataaatcc	ctgtccatga	aaatgcaaat	tatccaggtg	agatatattt	gacatcacc	7560
ctatgggat	tggctagttt	tgcoctttat	aagcaaatcc	atttcagcct	gaatgtctgc	7620
ctatatattc	tctgtctttt	gtattctctc	ttgaaccgtg	taaaaatccc	tgtggcactc	7680

<210> 379

<211> 2679

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002073

<400> 379

gagaccagga	ccatgggatg	tcggcaaaagc	tcagaggaaa	aagaagcagc	ccggcggtcc	60
cggagaattg	acgcgcaact	gcgctcagag	agccagcgcc	aacgcgcgca	aatcaagctg	120
ctctctgtcg	gcaccagcaa	ctcaggcaag	agcaccatcg	tcaaacatcat	gaagactcat	180
cacagcggcg	gcttcaactc	ggaggcctgc	aaggagtaca	agccctcagat	catctacaat	240
gccatcgact	cgctgaccgc	catcatccgg	gccctggccg	ccctcaggat	cgacttccac	300
aaccgcgacc	gcgcctacga	cgctgtgcag	ctctttgcgc	tgaacggccc	cgctgagagc	360
aaggcgagga	tcacacccca	gctgctgggt	gtcatgcgac	ggctctgggc	cgacccaggc	420
gcacaggcct	gcttcagccg	ctccagcgag	taccacctgc	aggacaacgc	ggcctactac	480
ctgaacgacc	tggagcgcat	cgccgcagct	gactatatcc	ccactctcga	ggacatcctg	540
cgctcccgcc	acatgaccac	gggcattgtg	gagaaacagt	tcaccttcaa	ggagctcacc	600
ttcaagatgg	tggagctggg	ggggcagagg	tcagagcgca	aaaagtggat	ccactgcttc	660
gagggcgctc	cagccatcat	ctctctgtgtg	gagctcagcg	gctacgacct	gaaactctac	720

```

gaggataacc agacaagtgc gatggcagag agcttgccgc tctttgactc catctgcaac 780
aacaactgggt tcatcaaac ctoactcacc ctctctctga acaagaagga cctgctggca 840
gagaagatcc gccgatcccc gctcaccatc tgctttcccg agtacaaggg ccagaacacg 900
tacgaggagg ccgctgtcta acctcagccg cagtttgaag ccttgaaacc caacaaggag 960
accaaggaga tctactccca ctcaactcgc gccaccgaca ccagtaacat ccagtttgtc 1020
tcgacgcggg tgcacgacgt catcatacag aacaatctca agtacattgg cctttgctga 1080
ggagctgggc ccggggcgcc tgccataggt gaaacccacg ggggtgcatg ccccaacggc 1140
tgcagagag ccccaatcca ggggcagaaa acagggggcc taaagaatgt ccccccccc 1200
ttggcctctg cctccttggc cccacatttc tgcaaacata aatatttacg gatagattgc 1260
taggtagata gacacacaca catgcacaca cacacatctg gagatggcaa aatcctctaa 1320
aatgtcgagg tctcttgaag acttgagaag ctgtcacaag gtcactacaa gcccaacctg 1380
cccttcaact ttgccttcc tgaattggccc cactccaact gggggtctgc attggattgt 1440
tagggatagg cagcagggct gaggcaaggt agggcaactg caacctctgc acctggagg 1500
gggcggcgct gctgcccgag ctctggccta gggacattgc cgtgaccaa gagggaggag 1560
cagtgcaagg tctgtgcacc ttccctgctg gactgcacac agctgctcag caccatttca 1620
ctctggacct gggacottag gagccgggtg acagcactaa ccagacccctc agccactcac 1680
agctcttttt aaaaaaacgc tcaaaaatat gcagcaaaaa ccaatacaac aaaaacgggt 1740
gccagtatta ttccaaacta gggcagctgg gattccaagct tttctctac tagtctgatg 1800
ttttataaat caaaacctgg ttttctctct ctggcatttt ttttgtttt ttgttttttg 1860
gttttttttt tttttttggc caaatctcgt ggtgtttcgc agaaaaaaat ccagaaaaat 1920
tcaaatgcag ttgagtttct ttttttaaat gcagattttc aaaaatatatt ttttttcagg 1980
tggctctttt tgtgtctggc ttgtcagtg taaagtgtt tatctggagc atctgtctct 2040
ctgtccaaa gaaatttttg agtgagtggc agtctcgcc cagcctcgcg gacacgtgtg 2100
tgtacataag cctctgcagt gtctcttgt taatggtggg gttttctgct ttgtttttat 2160
ttaagaaat aaacacgaca tatttaaaag aggttctttc acctggggag caatgaacaa 2220
tagctaaagt tcttggtatt taaagagtaa attatttgtt gctttgctga gtgaaggaa 2280
gggagcaagg ggtggcgccc ctgtgtccag cgtgccccc cctcgagact ggctggaaat 2340
gctctgaact ctgtgaaggc acagccacgc ttgtggcctg agggaggccc gctgggaacc 2400
ctgatctggg cctctctgtc ccagggccca tggggcaact gttgaaagg acgttgccca 2460
agggccgtgt gtaatacga actgcgccat tgtgcagtag ggcactggcg gagccttgc 2520
cagatctccc tccctctctc tgtacggggg ctgtgtgtcc gaggtcagtg tgccgaatca 2580
cagcaagga gtaagaaaga aataaaatcc aagaagaagg tgggattgg atgaagatca 2640
aaggttgtct actttaagaa aataaaatcc cctgaattg 2679

```

<210> 380

<211> 3337

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000436

<400> 380

```

gtcagacctc tagcccgccc gggtttctct cgcagtcgcy caccgacgct caaacgcgcy 60
ctccaaaccc cagcctctct ctcgctcacc gcccgaaagt ggcggtctct aactcctct 120
ctccggggtt tcggctctgc gccctgcgcc ggggaatctgg ggcaacctgg tacaaggat 180
gtgtttgttc ctttccacc agtgctcacc gccataccaa gttttataca gtaacagtag 240
aagctgtaaa agacatcccc gatggtgcc cggttttggt tgggtgtttt ggggtattgt 300
gaattccaga gaatcttata gatgctttac gataaaactgt agtaaaagga ctaactcgca 360
tcagacacaa tgcagggggt gacaattttg gtttgggggt tttgcttcgg tcgaagcaga 420
taaaacgcac ggtctcttca tatgtgggag aaaaatgcaga atttgacga cagtacttat 480
ctggtgaaat agaagtggag ctgacacacc agggcacact tgcagagagg atccgtgcag 540
gcggggctgg agttcccgca ttttacaccc caacagggtg atttgacga cgtacttat 600
gaggtacgct catcaaatat acaaaagatg gcaagtgttg tgggacccgt cattgccagt aagccaagag 660
aggtgagggg gttcaatggt cagcaccttta ttttgaggag agcaattaca agggattttg 720
ctttggtgaa agcctggaag cgggaccgag caggaacacat gattttcagg aaaaagtcaa 780
gaaatttcaa ctttgcaact tgcaaaagctg gctccagaag acatccatat tctctagatt ttgaaagaa 840
tttgtgatat tgggacattt tatgagaaaa gaattgagcg tttatcaatc cggaaaggag 900
gccattataa gggagaaaaa tatgagaaaa gctaaaacct gagatgacgt aagggaacga atcatcaga 1020
gagatggggg agccaaatct gctaaaacct gatggcatgt atgtcaattt gggcatagga atccctctcc 1080

```

```

tggccagcaa ttttatcagc ccaaatataa ctgttcacat tcaaaagtga aatggaggttc 1140
tgggtttggg tccatatcca cgacaacatg aagctgatgc agatctcatc aatgcaggga 1200
aggaaacagt tactattctt ccaggagcct ctctttcttc cagcgatgaa tcattttgcaa 1260
tgatttaggg tggacacgtc gatctgacaa tgcctaggagc gatgcaggtt tccaatatgt 1320
tgacctggcg taactggatg atacctggga agatggtgaa aggaatggga ggtgctatgg 1380
atttagtgtc cagtgcgaaa accaaagtgg tggtcaccat ggagcattct gcaaaaggaa 1440
atgcacataa aatcatggag aatgtatcat taccattgac tggaaagcaa tgtgtcaacc 1500
gcattattac tgaaaaggct gtgtttgatg tggacaagaa gaaaggggtg actctgattg 1560
agctctggga aggcctgaca gtggatgacg tacaagagag tactgggtgt gattttgcag 1620
tttcacaaa accatgagca atgcagcaga gttcaaattg aaatatggat attgtacca 1680
ggctgcgtgt ttttcatttt aaacacacaa gatttaattg aaaggacatc aataatcata 1740
attgtgtatt taacagggtg ttttttatta gtttctgtgt ttttcagact ttatgcagcc 1800
atataaactg ttctcttagg atgctgtgac attttaataa aaagcaaaag gacgattatt 1860
aattatactt ttgtttaagg ctgagaaggt gtgttttata ataggttaatt atattgaatg 1920
cattttcaat gaatatggta tgtatgctaa attatatgaa cctttcccca agaagggccc 1980
tagaaaatga tgggcttttc ctcttaataa ttaattatta gtccgaaag aagataaaca 2040
tatgtgattt ttgtggttag gagagttgct gtcagtattg tttttcttc agcctcctct 2100
gaactttctt ttggggcttc agattttatg attacatctt gtccccctag aacatcccc 2160
ttcctcccat acgtctttta acagatggc caagaaggca ctaggaaatg cctctgtggt 2220
gggagggcag ggagaaataa ctagtcca ccaactatct atctatgctt tgcaaaagct 2280
aaggctgatt ataggagag ggctagaac ctactgatt ctctcagtt ttctctattt 2340
aaaaacagccc agtattcctt tgtatcctca agggctcgtg agaatactc ttgtattgaa 2400
accctgtggg ctactgttac tgaacctctc gaagggcgtg gaaggaattt 2460
accatgaatc cttagtgaca atgacagcag agttaaaaaa taaaggtgt ttactttca 2520
ggctcttggt ttggttcaga ggagatttta atatttgat gacacttcta cagaacaacg 2580
gtttttcttc tgcgaaggct acttctttta acgaagtgc ttaattcag ccttatccaa 2640
cttggaaaaa taatgttggg caagcttagg attaggtgtt gaaaataact tccttagagg tttagtgtca 2700
gggaataaac atggtgggta gattaggttt gaaaaaaact tccttagagg tttatttctt 2760
caatacctga cagggggccca tgggaatgac ttcagaagca tcccggataa tagatgggtg 2820
aaaagtctag gcacctgaa gaacaggtga gacagctggc ctctggacag aggtaggcat 2880
agtcagctac gatatacat tctcttggtc tcaaatatac aaacttatc atgttttag 2940
gtgatgatgg tcattgaaac tcacttcttt tcaggtgtag ctacaattgt tttagtaca 3000
atattagaga aaggacaggc tttttatgag taacacacac catatataaa acagccttct 3060
tggctgacca catggtttaa tgcatacctt cccagtactg gggggaaaat gaccttctt 3120
agaattgtga agtccatgat agtaatatat tgatatgatt ttgaaaagaa ttgttgatag 3180
ttacattctc aaacttatca ttccagtagt catctttaag ataattgcat aaggaataca 3240
tgactttata ttcttgatta aagagtgtca tacatgttaa gaaatgcatt aaggaataca 3300
ataaatattc taactgatg aaaaaaaata aaaaaaaa 3337

```

<210> 381
<211> 1020
<212> DNA
<213> Homo sapiens

<300>
<308> Contig55377_RC

```

<400> 381
tagttttcaa aaaaattattg atctggatca gatgggctga taaggcccaa acactataa 60
accaaagctg caggctctcta aggtcttcag agaagctaca caaaatacta tgaatttcatt 120
agcatccaca tatgtatcat ttccacaac attttacaat tctctgtgtt cacagggaag 180
tgaaactaat tataatgaaa tatcagtaat gtatcttaaa taagtacaat gttgctgaaa 240
aaaatttaag ttaaaatgca ctacataata gtgcagaaaac aaagtgcagg ttgcaatttt 300
aagacatgaa aacacatgctg aatcagtagt gtttggagtg atacattttg cccatttttg 360
gccagtccta gaacagctact caccacaagc tgcactgcaa atatgaacta atgaggtcaa 420
aacactgaaag accaaataga accactctag atgtctgttc acaatgagat acctctcccc 480
gttctggaaa cgcgactctg ctgtagttcc tatctcctg tggcacctat cctcgctaca 540
tgactcataa ccagatgttc tccatcatcc agcoctcagg acaagtggtg ctacagtagt 600
cacaacctga gaaagcagctg ctggttaccg atttgtatc ctgtatatca tgcagttggg 660
tagccaggat ttctctggtt gtcttgaata tatttacttc taaaagact gttagtcttg 720
ttgggaggtg agcagaggaa tgtttacttg ctggagaaca tagtaacaaa catcactctt 780

```

ccccctcat aaacagaaag ccattgtct gtgccgtgaa aagaggacta cctgtcagg 840
aagattttac ttacctctac aagaaaaagg actcgtatta aaatcgacct gtggttacca 900
cccttagaaa tgcgtttttc agatattcct attcctagtc tttagaggcag ccatggaata 960
tatagtctcc tctgaccttg gggcaggaa gaaagggaat tctattcata aaagggaatt 1020

<210> 382
<211> 668
<212> DNA
<213> Homo sapiens

<300>
<308> Contig35251_RC

<400> 382
ctatacagta ttgttcagct gtttttattc ttacgtttat catcttatat ttggctctca 60
atttcaccca taaattattg atggcttatt atacgccaaag tacctctgctg tccgcttttt 120
gtggatcccc tcttcggagc gagaagtcac ttagaatttt acccatcttt cccactttgg 180
gaatcggaaag ggttcatagc taagttttgt tattataaag cacgccggct cttctctatc 240
cctctctcca ccaaccccat ttcaagggtg gtccctccca gcggacgaca gtgatgaatt 300
cctcttgggg acagccctcc cgcaactaac ggtggctcca tcgacgtctt ggtctccttg 360
ctgctcgctg ggaattccgg tgggacgttt tcctaaagttt ggattccacg gcacattctc 420
ccctcggaaa ggacacaacc gctcggcgctg tcagctcagc tggactcacg ccggtgcggg 480
gcgcgagctc cggttctcgc tggggggctg cggagggctg ggcgattctc tcagcgtcgc 540
gaggcgaccg gggcagcagc aaacttctcc gcttccctc ttccgggacg gcacactctt 600
gtgcccttga gcactccgca ggaacttccc ccggcctcgg tctcgtccg aaataaccga 660
gccagctc 668

<210> 383
<211> 430
<212> DNA
<213> Homo sapiens

<300>
<308> Contig25991

<400> 383
agagggtcg gcgtttgaag aggtggaact cctagggctt ttttagagat gctgatttag 60
aagaatacaa atcatggctg aaaaatagtg attaacatcc actactggga ggactagctt 120
ggcagactct gcactttttg attctaagtg tactgagatt tccaaggaaa acttacttat 180
tggtattact tcatatgtag aagaagagat gcctcagatt gaaacaagag tgatattgtg 240
tcaaggagct ggaaaacaag aagaacttat aaaaagccta aaggacatta aagtgggctt 300
tgtaaagatg gagtcagtgg aagaatttga aggtttgtgt tctccggaat ttgaaaatgt 360
attgtagtc accgacttcc aggattctgt ctttaatgac ctctacaagg ctgattgtga 420
agttattgga 430

<210> 384
<211> 2212
<212> DNA
<213> Homo Sapiens

<300>
<308> NM_003875

<400> 384
tgccgggtcc tccctgacca ggctcctctc tcaacctcag ccgcgggcgc cgacccttcc 60
ggcaccctcc cgcccgtct cgtactgtcg ccgtcacccg ccgaggctcg gccctggccc 120
cgatggctct gtgcaacgga gactccaagc tggagaatgc tggaggagac cttaaggatg 180
gcaccaccca ctatgaagga gctgtgttca ttctggatgc tgggtgctcag taagggaag 240
tcataagacc aagagtggag gaactgttcc tgcagtcgta aattttcccc ttggaacac 300
cagcatttgc tataaaggaa caaggattcc gtgctattat catctctgga ggacctaat 360

ctgtgtatgc	tgaagatgct	ccttggtttg	atccagcaat	attcactatt	ggcaagcctg	420
ttcttggaat	ttgctatggt	atgcagatga	tgaataaaggt	atttgagggt	actgtgcaca	480
aaaaaagtgt	cagagaagat	ggagttttca	acattagtgt	ggataataca	tgttccattat	540
tcaggggcct	tcagaaggaa	gaagttgttt	tgcttacaca	tggagatagt	gtagacaaag	600
tagctgatgg	attcaaggtt	gtggcacgtt	ctggaaacat	agtagcagcg	atagcaaatg	660
aatctaaaaa	gttatatgga	gcacagttcc	accctgaagt	tggccttaca	gaaaatggaa	720
aagtaactact	gaagaatttc	ctttatgata	tagctggatg	cagtgggaacc	ttcaccgtgc	780
agaacagaga	atttgatgtg	attcgagaga	tcaaaagagag	agtagggcag	tcaaaagtgtt	840
tggttttact	cagtggttga	gtagactcaa	cagttttgtac	agctttgtcta	aatcgtgctt	900
tgaaccaaga	acaagtcatt	gctgtgcaca	ttgataatgg	ctttatgaga	aaacgagaaa	960
gccagtctgt	tgaagaggcc	ctcaaaaagc	ttggaattca	ggtcaaagtg	ataaatgctg	1020
ctcattcttt	ctacaatgga	acaacaaccc	taccaatctc	agatgaagat	agacccccac	1080
tgaaaagaatt	tagcaaaacg	ttaaatatga	ccaagaatcc	tgaagagaaa	agaaaaatca	1140
ttggggatgc	ttttgttaag	attgccaatg	aagtaattgg	agaaatgaac	ttgaaaccag	1200
aggaggtttt	ccttgcccaa	ggtactttac	ggcctgatct	aattgaaagt	gcattccctg	1260
ttgcaagtgg	caaaagctgaa	ctcatcaaaa	cccatcaca	tgacacagag	ctcatcagaa	1320
agtgtagaga	ggagggaaaa	gtaatagaa	ctctgaaaga	ttttcataaa	gatgaagtga	1380
gaattttggg	cagagaactt	ggacttccag	aagagttagt	ttccaggcat	ccatttccag	1440
gtcctggcct	ggcaatcaga	gtaatatgtg	ctgaagaacc	ttatatgttg	aaggactttc	1500
ctgaaaccaa	caatattttg	aaaatagtag	ctgatttttc	tgcaagtgtt	aaaaagccac	1560
ttaacctatt	acagagagtc	aaagcctgca	caacagaaga	ggatcaggag	aaagctgatc	1620
aaattaccag	tctgcattca	ctgaatgcct	tcttgctcgc	aattaaactc	gtagggtgtc	1680
agggtgactg	tgttctctac	agttacgtgt	gtggaatctc	cagtaaagat	gaactgtact	1740
gggaatcaact	tatttttctg	gctaggctta	tacctogcat	gtgtcacaac	gttaacagag	1800
ttgtttatat	atttggccca	ccagtttaag	aaacctctac	agatgtttac	cccactttct	1860
tgaacacagg	ggtgctcagt	actttacgcc	aagctgattt	tgaggcccat	aacattctca	1920
gggagctcgg	gtatgctggg	aaaactcagc	agatgccgtg	gattttgaca	ccattacatt	1980
ttgatcggga	cccaactcaa	aagcagcctt	catgccagag	atctgtgggt	attcgaaact	2040
ttattactag	tgaactctatg	actggtatgc	ctgcaacaac	tgccaatgag	atccctgtag	2100
aggtggtatt	aaagatggtc	actgagatta	agaagattcc	tggtatttct	cgaattatgt	2160
atgacttaac	atcaaaagccc	ccaggaacta	ctgagtgga	gtaataaact	tc	2212

<210> 385

<211> 2150

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006101

<400> 385

ctcgagccac	gaaggccccc	ctgtcctgtc	tagcagatgc	ttgcacgggt	tacagaaatt	60
cggtccctgg	gtgctgtcag	gaaactggaa	aaaaggtcat	aagctagaag	cgcagttcag	120
tttccagcgg	tggtgtctgc	cgctctccca	tgcaggagtt	aagatcccag	gatgtataat	180
acaaagccct	ctataccctc	caaaccaaa	gaataaccac	ctttggaaag	ttgagtataa	240
acaaaccagc	atctgaagaa	aaagtctcgc	tatttggcaa	aagaactagt	ggacatggat	300
ccgggaatag	tcaacttggt	atatattcca	gttctgagaa	aatcaaggag	ccgagaccac	360
ttaatgacaa	agcattcact	cagcagtgta	ttcgacaact	ctgtgagtat	cttagacaaa	420
atggtttatg	acataatgtg	tccatgaaat	ctctacaagc	tccctctgtt	aaagacttcc	480
tgaagtcttt	ccactttctt	tatggcttcc	tgtgccctcc	atcgaacttt	cctgacacaa	540
agtttgaaga	agaggttcca	agaatcttta	aagaccttgg	gtatcctttt	gcactatcca	600
aaagctccat	gtacacagtg	ggggctctcc	atacatggcc	tcacattgtg	cgagccttag	660
tttggctaat	agactgcac	aagatacata	ctgccatgaa	agaaagctca	cctttatttg	720
atgatgggca	gccttgggga	gaagaaactg	aagatggaat	tatgcataat	aagttgtttt	780
tgactacac	catataaatgc	tatgagagtt	tattagtggt	tgccgacagc	tttgatgaga	840
tgaattcaga	gctgcagctg	aaactgaagg	atttatttaa	ttgggatgtc	tttaagctgg	900
aatcatatga	agcaaaaaac	agagcattga	atgaacagat	tgcgaagattg	gaacaagaaa	960
gagaaaaaga	accgaatcgt	ctagagtcgt	tgaaaaaact	gaaggtctcc	ttacaaggag	1020
atgttcaaaa	gtatcaggca	tacatgagca	atttggagtc	tcattcagcc	attcttgacc	1080
agaaattaaa	tggtctcaat	gaggaaattg	ctagagtga	actagaattg	gaaacaataa	1140

aacaggagaa	cactogacta	cagaatatca	ttgacaacca	gaagtactca	gttcgagaca	1200
ttgagcgaat	aaatcatgaa	agaaatgaat	tgacgacag	tattaataaa	ttaaccaagg	1260
acctggaaag	tgaaccaacg	aagtgttgga	atgaggagtt	aaaatatgct	agaggccaag	1320
aagcgattga	aacacaatta	gcagagtgat	acaaattggc	tagaaaatta	aaacttattc	1380
ctaaaggtgc	tgagaattcc	aaaggttatg	actttgaaat	taagtttaat	cccgaggctg	1440
gtgccaactg	ccttgtcaaa	tacaggggctc	aggtttatgt	acctcttaag	gaactcctga	1500
atgaaaactga	agaagaat	aataaagccc	taataaaaaa	aatggggttg	gaggataact	1560
tagaacaatt	gaatgcaatg	ataacagaaa	gcaagagaag	tgtgagaact	ctgaaagaag	1620
aagttctaaa	cgctggatgat	ctttaccaac	aaaaaattaa	ggaagcagag	gaaggaggtg	1680
aaaaaatgca	gagtgcgctt	gagtcctttg	agaaacacaa	gcacctgcta	gaagtagact	1740
ttaaccaggg	gctcagtgaa	gctatgaatg	aattagatgc	tgttcacagg	gaataccaac	1800
tagttgtgca	aaccacgact	gaagaaagac	gaaaagtggg	aaataacttg	actgaactgt	1860
tagagatggt	tgctacacat	gttgggtctg	tagagaacaa	tcttgaggag	cagattgcta	1920
aagttgatag	agaatatgaa	gaatgcattg	cagaagatct	ctcgaaaaat	attaaagaga	1980
ttagagataa	gtatgagaag	aaagctactc	taattaagtc	tcttgagaaa	tgaagataaa	2040
atgtgatcat	tgtatatata	tccatagtg	ataaaattgt	ctcagtaaaa	aaaaaaaaaa	2100
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa		2160

<210> 386
 <211> 2830
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003882

<400> 386	ccaacgcgtc	cgctggggcc	agctccccg	agaggtgttc	ggatcctctg	ggctgctcgg	60
	tcgatgcctg	tgccactgac	gtccaggcat	gaggtggttc	ctgcccgtga	cgctggcagc	120
	agtgcacgca	gcagccgcga	gccacgcgtc	ggccacggcc	ctctctccag	ccctcagac	180
	catggacttt	actccagctc	cactggaggga	cactcctcca	cgcccgaagt	ctctgaagtg	240
	gcaatgtgag	tgcccgcgat	ccccaccocg	ctgcccgcgt	ggggtcagcc	tcatacaga	300
	tggtctgtgag	tgctgttaaga	tgctgcgtca	gcagcttggt	gacaactgca	cggaggtctg	360
	catctgtgac	ccccaccggg	gctctactgt	gggtcgtgct	ggggaccgcc	cgaggtatgc	420
	aatagagagt	tgctgcacagg	tggtcggtgt	gggtcgtgct	ctggatgggg	tgccgtacaa	480
	caacggccag	tccttccagc	ctaactgcga	gtacaactgc	acgtgcacgt	acggcgcggg	540
	gggtgcacac	ccaactgtgc	tcggagtgcg	cccccccggt	ctctggtgct	ccccaccocg	600
	cgctggtgac	atacctggcc	actgctgtga	gcagtgaggta	tgtaggaagc	acgccaaag	660
	gcacacgaag	acgcaccccc	gtgacacagg	agccttcgat	gctgtgggtg	aggtggaggc	720
	atggcacagg	aactgcatag	cctacacaag	ccccctggag	ctctgctcca	ccagctgcgg	780
	cctgggggtc	tcacactcga	tctccaatgt	taacgccagc	tgctggcctg	agcaagagag	840
	cgctcctgcg	aacttgcggc	catgcgatgt	ggacatccat	acactcaata	agggcaggga	900
	gaagtgtctg	gctgtgtacc	agccagaggc	atccatgaac	ttcacacttg	cgggctgcac	960
	cagcacacgc	tcctatcaac	ccaagtactg	tggaagtgtc	atggacaata	gtgctcgat	1020
	ccctacaag	tcctaagacta	atgcctgctt	cttcacagtg	cctgatgggc	ttgctttctc	1080
	cgccagaggtc	ctatagatta	cctaacctga	cttcacagtg	agctgtagga	atcccaatga	1140
	catctttgct	gaacttggat	taacccaatg	cctgtgaagc	attgccacct	agggcaggac	1200
	aaatcttggt	tcttggggac	taagttaccct	gatctggacc	cttggcctcc	attttctgct	1260
	aacttttcaac	caatgagcct	tgatggtgct	gctcaggccc	atgctatgat	ttttctctct	1320
	ctaaacattc	aaatgagccc	ctaaagaaaa	agaagtcctg	ctggagtctg	tgagctacac	1380
	gatatacttc	agcatctact	ccaagtcact	gctcagctgt	ctgaggtctg	ctctaatccc	1440
	ccaagcctga	tcacagcctt	cttttaatat	gctaatttc	ttcttttagat	cgaaaaccac	1500
	aagaaatgga	atcaggtaga	agatgaatat	atggaatttg	gaacaataga	ataatctatt	1560
	aaagctctctt	gggtccattc	tactgtaattg	ggttaattctg	acctaagcgc	acaaaacta	1620
	atttggagcg	tgccaaaggg	gcacctcaag	gtcatcaaac	atttggccag	tgagttgaat	1680
	tcctgatccc	aaatattgat	taatggaaa	ttgtatccat	taacctgggg	attgttgagg	1740
	agttgctctaa	ttttgatttt	acaactgtga	gggtacagat	taggtttgtc	ccagtcagaa	1800
	ttaaagttct	cttcaaccct	ctgttgatgg	gaaaagcccc	cagtttaaac	tcacagagaa	1860
	ataaaatttg	atgaacattc	agaaggacca	attgactctc	acactgaact	agctgctgac	1920
	gggaaaggctc	agcccatctc	ggccaggctc	ttccttgaat	cttctccctt	gtcctgcttg	1980
	tgccagggct	ttgggacggt	ggccaggctc				2040

ggttcataagg	aattggtaag	gcctctggac	tggcctgtct	ggccctgag	agtgggtgcc	2100
tggaacaactc	ctctactctt	acagagcctt	gagagaccca	gctgcagacc	atgccagacc	2160
caactgaatag	accaagacag	gttcaggtag	gggtgtgggt	caaaocaaaga	agtggtgtgcc	2220
cttggtagca	gcctgggggtg	acctctagag	ctggaggctg	tgggactcca	ggggcccccgc	2280
tgttcaggac	acatctattg	cagagactca	tttcacagcc	tttcgttctg	ctgaccacaa	2340
ggccactttt	ctggtaggaa	gatggaggtt	taccagtgtg	ttagaacacg	aaatagactt	2400
aataaaggtt	taagctgaa	gaggttgaa	ctaaaaggaa	aaggttggtg	ttaatgaata	2460
tcaggctatt	atttattgta	ttaggaaaat	ataatattta	ctgttagaat	tcctttattt	2520
agggcctttt	ctgtgccaga	cattgtctct	agtgctttgc	atgtattagc	tcactgaatc	2580
ttcacagcaa	tgttgagaag	ttccattat	tatttctgtt	cttacaatat	tgaacccgaa	2640
gctcatagag	gtgagaaac	tcaaccagag	tcacccagtt	ggtgactggg	aagattagga	2700
ttcagatcga	aattggactg	tcctttatac	ccatatttcc	ccctgttttt	taagacttcc	2760
aaatgtgtca	gaatgagaaa	acattgcaat	aaatggcttg	atttttttaa	aaaaaaaaaa	2820
aaaaaaaaaa						2830

<210> 387

<211> 2785

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003607

<400> 387

tcgcgaagcc	tcgcctctg	tgcgcgggac	tggaggagcc	tcgctgagcc	cagggcgcca	60
cgccagaggg	agaggggaag	cgggggaggg	ctggaaggga	gaggaaggga	gtggttgga	120
gcgggggtcg	cgcagcctct	agctctctca	gcgcgggaag	caccctcctc	ctgcgcgcgc	180
ggccgcctcc	ctcctcgctg	tggaagatg	cccttagccc	aggggtgtga	agaaagggga	240
gaagttagct	ccagagccgc	ccgcgcgcgc	gcgcagcct	ctgcctgttg	tgctcttga	300
attcagagaa	gcacccccc	gcctggggcg	tcggagacct	ccgggcggcg	gcgcgcgtgc	360
gggtgtgggg	agacccggct	ctctgcccg	gcggcgccgc	gcggctcgcc	ccacgagcga	420
ccacgcacat	ggagtggct	cgggcgccca	agtagccgct	tcctccggagc	cgggtccagc	480
gcgcgccgca	gccgccttc	caccccgcc	cgccgcgcgc	gtcaggccct	aggggtgaagc	540
cgggaggaaa	atgaagatgt	ttcacccgga	atccgttgaa	aatgagactg	actgcgaagc	600
cttaagaaaa	gaagacctc	gggaggagaa	acgaaaagcc	gcctccggcg	aagacttgcc	660
gtgctccgag	ccgaggggct	gcttcaggga	cctgcgcgcc	tcctttcccc	gctggagaaa	720
ttgcccgtga	tgcattatcc	aagtgtgtgt	tgggaggatt	tgcagcaaca	ttttgtggtt	780
tcctccccc	ttctatgcat	tcgtttttt	tcctcccttt	tcgtttttc	tcctcccg	840
gaagtgaatt	gctgatgcaa	atcggacttt	attcatat	gatgcaaccg	gattcgtttc	900
aggattacgt	gcgcaggtt	gaattttgaa	tgaaggagaa	gagttttttt	tttttttaa	960
agaagtgttg	actctctagt	tcgttgtact	tttaattatt	attttattta	aatatacgac	1020
tttaattgat	tcgtttaaaa	atgcattaa	tatatattt	atggttaatt	accctcaaaa	1080
tatatgtata	tggttgaat	tgaagacgcc	ttcagttaa	tgaggttact	ggtgtgtgtg	1140
atgttttaatt	cagaccacgc	attgcatgac	agttgtttga	ataacaagtg	gtttattttt	1200
aaaacactac	cttttaaaat	ttaggttcag	ataatagtaa	aagtcacatc	aatattttaa	1260
aggaataacca	tcggaatact	aagcaaacat	gtctggagaa	gtgcggttga	ggcagtttga	1320
gcagtttatt	gtgcgcggc	ccgctcagac	caatgggcag	tacttcagtg	tggaagacatt	1380
actgatatata	ctcatctgcc	tttatgatga	atgcaataat	tcctcattga	gaagagagaa	1440
gaacattctc	gaactacctag	aatgggctaa	acacatttact	tcataagtg	acacaaatgc	1500
attacataga	gaagactttg	aaatattaaa	ggtgatgtgt	cgaggagctt	ttggggaggt	1560
tgtcttagta	aagctaaaaa	atgcagataa	atgagtttgc	atgaaaaaat	tgaataaagc	1620
ggaatgtctg	aaaagagctg	agacagcatc	tttctgtgaa	gaagaggatg	tattagttaa	1680
tggaagacaa	aaatggatta	caaccttgca	ctatgctttc	caggatgaca	ataactata	1740
ctggtgttat	gaattattatg	ttggtgggga	tttgcttact	ctactcagca	aatttgaaga	1800
tagatttgct	gaattatagc	ctagatttta	ctggctgag	atggtgtgat	caattgactc	1860
agttcatcag	ctacattatg	tacacagaga	cattaaacct	gacaaatata	tgatgtgat	1920
gaatggacat	attcgggtag	cagatttttg	tcctgtctgc	aagctgatgc	agaatgggac	1980
ggttcagtc	tcagtggctg	taggaactcc	agattatata	tcctctga	tcctcgaagc	2040
catgggaagt	ggaaaaggga	gatattggac	tgaatgtgac	tggtgtgctt	ttgggggtctg	2100
tatgtatgaa	atgctttacg	gagaaacacc	attttatgca	gaatcgctgt	tggagacata	2160
cgaaaaaatc	atgaaacaca	aagagaaggt	tcagtttcca	gcccaagtga	ctgatgtgtc	2220

tgaaaatgct	aaggaatctta	tccgaaggct	catttgtagc	agagaacatc	gacttggtca	2280
aaatggaaata	gaagactctta	agaaacaccc	atttttcagt	ggaattgatt	gggataatat	2340
tcggaactctg	gaagcacctt	atattccaga	agtttagtagc	ccaacagata	catcgaattt	2400
tgatgtagat	gatgatgtgt	taaaaaatc	tgaacgatg	ccccaccac	cacatactgc	2460
attttctggc	cacatctgc	catttgttg	ttttacatat	actagttagt	gtgtactctt	2520
tgatcgagc	tggttgaag	taccggctgg	tccacctca	ctggtctgtg	atgttaatgt	2580
tcagaggact	ctagacaaca	acttagcaac	tgaagcttat	gaaagagaa	ttaaagcgct	2640
tgagcaagaa	aaacttgaa	tcagtagaaa	acttcaagag	tcaacacaga	ctgtccaagc	2700
ctctgcagat	tcaactgttg	atgttccact	aacagcaagc	aaagatttag	aaataaaaaa	2760
cttaaaagaa	gaaatttgaa	aaaaa				2785

<210> 388

<211> 1912

<212> DNA

<213> Homo sapiens

<300>

<308> AF073519

<400> 388

ccggcggttg	ttgtcgggcc	tccagcgggc	ggggcggtg	gcggagcaga	gcggaggcgc	60
agccggcggtg	agggccacag	agggtccagc	cttcccggtc	agcgggtggtg	acggtatccc	120
agatggccag	agaacccgttg	cttttccgag	ttgctctctc	tccaggctcc	gttggtgggc	180
ggcatggccc	gtggaaatca	acgagaactt	gcccgccaga	aaaacatgaa	gaaaaccacg	240
gaaattagca	agggaaagag	gaaagaggat	agcttgactg	ctctccagag	aaagcagagt	300
tcaggaggcc	agaaatctga	gagcaagatg	tcagctgggc	cacacctccc	tcgtaaggct	360
ccaaaggaga	atccttgctt	tcctcttcca	gctgctgggtg	gctccaggtg	taacttgggt	420
ttggcagca	taactcctat	ctctgccttt	gtcttctgtg	ttcttctttc	tgtcttcttc	480
ccctctcttt	atgaggactt	ttgctgttgg	atttaggttc	cattctaaac	taggatgato	540
tcatttggaa	atccttaatt	tcactctaca	aaactgtttt	cccaaatagg	tcacattcac	600
gcataatcaga	tgacacagatg	tatcattttg	gggtccacca	ttcaaccacc	tacaaggagt	660
tttttaacaa	aaaataggaa	acttagatgt	aacttagcac	tttttttttt	tttttttgag	720
atggagatctc	actctgtcac	cagactggag	tgcaagtggc	ccatctcagc	tcocatgcaac	780
ctctgcctcc	tggtgtcaag	cagttctctt	gcctcagcct	ccctgggtagc	tcggattaca	840
ggcagcgctg	gccacaccca	ggtatattat	ttattttttt	tttgagacag	agctctcgac	900
tggttggcag	ctgggactgc	agtgcggtga	tcctctgtca	ctgcaacctc	gcctcccgcg	960
gttcaagcga	ttctccagcc	tcagcttctc	gagtagatgg	gattacaggc	gcctgccacc	1020
acgcccgagct	aatttttttg	tattcttagt	agagatgggg	tttcaccatg	ttggccaggc	1080
tggtctccat	ctccagacct	cgtagttcac	ccgcctcggc	ctcccaaatg	gctgggatta	1140
caggcgtgag	tcacagcccc	cgcccataat	ttagcacttt	aaaaataaat	agccattcac	1200
ggccaggcgt	ggtggctcat	gcctgtaatc	tgagcacttt	gggagaccaa	ggcgggtaga	1260
tcctctgtgc	ccaggagttc	aaagaccagc	tgggcaacat	ggcgaaaccc	catcttctact	1320
aaaaatacaa	aaattagctg	gggcgagggg	ataggccgag	ttccgggtgt	aaagggggcca	1380
ttaggagagag	cagagcgagg	cagctgatct	tcgggattgg	gggccttgcc	cggaagctgg	1440
acctcacgga	gatgaaacgg	aaagtgcacg	aggatatgat	ctccatacac	aaccttctca	1500
ttcactgtgc	ctctgctgca	gtcactccat	ttatcttaaa	gaaattggag	agcatatgaa	1560
gattggacat	cacatgtgaa	tgcatgatat	gaagagcctg	gttacagttt	ctactgttct	1620
ctcgaaagtaa	ataggccagc	aaaggtataa	gagactcttt	gaatggacat	aaaaattctg	1680
cttgtttaaga	acaaagttgag	ctctggtaac	tgactctta	taaaaaat	taaaaaat	1740
tggaagactct	gaaatggagt	ctctcgccgc	tggtgtgccc	ttaatgcctg	tgacagttgg	1800
ccctctgtaa	tattgtgata	attgtaata	atgtcaaac	ccattttcta	gcaagtatta	1860
ataattaagg	gaagtatgtc	tgaattggca	aaaaaaaaa	aaaaaaaaa	aa	1912

<210> 389

<211> 1266

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000849

<400> 389

```

ctcgggaagcc  cgtcaccatg  tctgtcgagt  cgtctatggt  tctcgggtac  tgggatattc  60
gtgggctggc  gcacgccatc  cgcctgctcc  tggagttcac  ggatacctct  tatgaggaga  120
aacggctaac  gtgcggggaa  gctcctgact  atgatcgaag  ccaatggctg  gatgtgaaat  180
tcaagctaga  cctggacttt  cctaactctg  cctacctcct  ggatgggaag  aacaagatca  240
cccagagcaa  tgcocatctg  cgtacatcgc  ctgcgaagca  caacatgtgt  ggtgagactg  300
aagaagaaaa  gattcagagt  gacatcatag  agaaccgaat  aatggatttc  cgcacaacac  360
tgataaggct  ctgttacagc  tctgaccacg  aaaaactgaa  gcctcagtag  ttggaagagc  420
taacctggaca  actgaaacaa  ttctccatgt  ttctgtgtaa  attctcatgt  ttgcccgggg  480
aaaagctcac  ctttgtggat  ttctcaacct  atgatattct  ggatcagaac  cgtatatttg  540
accccgaagt  cctggatgag  ttcccaaac  tgaaggcttt  catgtgcgtt  ttgaggctt  600
tggagaaaa  cgtgccttac  ttacagtctg  atcagttctg  caagatgccc  atcaacaaca  660
agatggacca  gtggggcaac  aagcctgtat  gctgagcag  aggcagactt  gcagagcttg  720
ttttgttcca  tccctgcctg  aaggggtcag  cgtcttctgt  ttgctctttt  caatgaatag  780
cactatgtt  actggtgtcc  agctgagttt  ctcttgggta  taaaggctaa  aagggaaaaa  840
ggatatgtg  agaatactca  agatatgaat  tgaatcgctg  cgatactgtg  gcatttccct  900
actccccaac  tgaagtccaag  ggctgtaggt  tcatgcccaa  gccctgagag  tgggtactag  960
aaaaaacgag  attgcacagt  tggagagagc  aggtgtgtta  aatggactgg  agtccctgtg  1020
aagactgggt  gaggaataca  caagtaaac  tgtggtactg  atggacttaa  ccggagttcg  1080
gaaacccgtc  tgtgtacaca  tgggagttaa  gtgtgataaa  ggcagtattt  cagactgggt  1140
ggctagccaa  taagtttggc  aattgcttat  tgaacctcat  taaaaataat  agagccccc  1200
ttgacactat  tcaactaaa  taatctggaa  ttaagggccc  aacattaaac  acaagctgtg  1260
attgat
1266

```

<210> 390

<211> 517

<212> DNA

<213> Homo sapiens

<300>

<308> Contig32185_RC

<400> 390

```

ttaaattagag  tttgtagtat  acacttaact  tctgttaact  tgcagtgatg  aatctctatg  60
taaacatttca  gaaaaagagc  gaatactgtg  tcatgactta  tgaactataa  attttgtccc  120
tgattactag  tcaagaacag  ttttatacat  ttattaaaca  ctactttcct  ccaacttaaca  180
acagctcttta  aataacaaaa  taaattatca  gcattttcaa  ccacaaattt  aagaattaaa  240
gtttaaaaaaa  ctggttagtt  tcataaatgc  aagcagagta  ataaatttca  tatatgttaa  300
tttgcagttta  ttatgttctt  tgagtgaagt  aacacaaaga  tttttccctt  ccccttttgg  360
caagtttaca  aaattcatga  aaagtcttca  cctttttctt  ttgtgtctaa  ttatattttg  420
gcattctcagc  tgaacacaac  agacaatagc  tggggttatt  agcagacat  gataaaagaa  480
aagggtatga  ccaactcagc  gatcatgtat  acctaaa
517

```

<210> 391

<211> 1266

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016577

<400> 391

```

gcgcgcgcgtc  cctgcgcccc  gctccgcgcg  cgcctctctc  ccagcgcgcg  gctcggggcg  60
taaaagagcgc  gcccttcgca  ccgcagccag  tgcgcagcgc  agcagcccaag  ccccaagcctt  120
cctcgcgcctc  cgcctcctct  tctcctcctc  cctccgcgcg  cgggaacgca  gccgaagcgc  180
ggaccgggac  gcagctgggg  agtcagggac  gcgcgcagcc  agcccttccc  cctccggctc  240
ccgcacgcgc  ggcgcgcctc  cctcgcctcc  ctactctccc  ctccctctct  ctctcgtttt  300
tctctcctct  cctctcctcc  ccccgctcgc  cagcaccatg  tcgcaggggg  gagatttttg  360
gaaatccactg  agaaaaatca  agttggtgtt  cttgggggag  cagagcgtcg  ggaagacgtc  420
tctgattacg  aggttactgt  acgacagctt  cgacaacaca  taccaggcaa  ccattgggat  480
tgacttcttg  tcaaaaacca  tgtacttgya  ggacgcgacg  gtgcgactgc  agctctggga  540

```

```

cacagctggt caggagaggt tccgcagcct gatccccago tacatccggg actccacggt 600
ggctgtgggt gtgtacgaca tcacaaatct caactccttc caacagacct ctaagtggat 660
cgacgacgtc aggcacagaga ggggcagtga tgttatcatc atgctgtggg gcaacaagac 720
ggacctggct gataagaggg agataacct cgaggagggg gagcagcgcg ccaagaacct 780
gagcgctcatg ttcatgtgaga ccagtgcgaa gactggctac aacgtgaagc agctttttcg 840
acgtgtggcg tcggctctac ccggaatgga gaatgtccag gagaaaagca aagaagggat 900
gatcgacatc aaagctggaca aacccaggga gcccccggcg agcgaggggc gctgctcctg 960
ctaattgcaga gccgacctgt ggcttcccat gacactcctt gcttgtgtgt ttgcttccct 1020
ttggctagct tccaaagggg ggagggaacc gagttatcaa gatggggagg tttttctttt 1080
ctctctgtct ttaggagtag ggtgggatgg ggaggaggcg tgggcatcag ggatcacatc 1140
actcttaacg cgtgttactt aaacaactat tttttggttt ggttgaata tattgtactt 1200
tattaagatt gccaaaaact gttaaaattt aaaaaaattt taaatcatgt gtatacaatt 1260
ttttgc

```

<210> 392
 <211> 472
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig48328_RC

```

<400> 392
gtacttttat tgaaggagta catttaaaaa aatacacaga cattttacca ttacaggtt 60
gcagatatag atgctctaaa agagtccact ctattttgtt gttotatgat aactctggc 120
cotgatatca caaacattcc agtctgtttg atatcggttt agaaaggggg gcatggggagc 180
atgacctgca atatattcag cgaacagaaa gacaaattgt tcaattataa atttttttat 240
cttctgtaca ttattgcatc agggagccac aaaattatgt agcatcatta caaatgaaaa 300
caggttaaaa atgaagaaga tacttatata gaaatacatg gatccattgt cttctgtcag 360
aatgcacaag aggtgcaaaa atgtgcaatt taggaagctc tttttctgtt tgtatcagtt 420
tgcttagcaa cacaaaccag tgagggaagct acaaaaatag ttaacacaaa at 472

```

<210> 393
 <211> 431
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig46223_RC

```

<400> 393
cttgtccaag atcattttatt gttattatta tttttctctc ttaattggat tcagcggttc 60
tgctctctca cacacatcct ctctggtggg ggattcaggt ttgtgcaaa gcaaaaggag 120
cttgtaccac cotgggcccc cggccagcag caaggcgccg ttgtatcgga gccctctctc 180
cctgtccggg ctgcctcgcc ccagtcacct ctctcgacc ctgagtcagg cctcccgagc 240
ctctgcctgc accctgagca tccccaggac ggtctggagg gccaggtctt ccgtgatgca 300
gaagtgctcg agatggttca tgaatatgcc cttctgttcc cttctgcta aggcoccgcg 360
acgcctccct cctgctgtcc cacacagggc cctgtctggg tcacacacca aacagcctca 420
aggaaaaatt c
431

```

<210> 394
 <211> 1728
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_015984

```

<400> 394
cggggaagtag gcgggggtgac gtgtggttga cgagctcggc ggcggggttg ctgagatctg 60
tgcccgccggc cagctgtgtgc gggggggcag tgagagcgag aggtggatcg gggcggtgtg 120

```

```

tggccagggc catgacgggc aagccgggg agtgggtgct catgaaagc gacccgggg 180
tcttcaccga gctcattaaa ggattcgggt gccaggagac ccaagtagaa gaaatagga 240
gtttgagagcc tgagatcttt gaaaatttaa agccagtcca tgggttaatt tttctttcca 300
agtggcagcc aggaagaagaa ccagcaggct ctgtggttca ggactcccga cttgacaaga 360
tattttttgc taagcaggta attaataatg cttgtgctac tcaagccata gtgagtgtgt 420
tactgaactg tacccaccag gatgtccatt taggcgagac attatcagag tttaaagaat 480
tttccaaaag tttttgtgca gctatgaaag gcttggcact gagcaattca gatgtgattc 540
gacaagtaca caacagtttc gccagacagc aaatggttga atttgatagc aagacatagc 600
caaaagaaga agatgctttt cactttgtca gttatgttcc tgttaatggg agactgtatg 660
aattagatgg attaagagaa ggaccgattg atttaggtgc atgcaattca gatgttggga 720
tcagtgctagt aaggcctgtc atagaaaaaa ggatacaaaa gtacagttaa ggtgaaattc 780
gatttaattt aatggccatt gtgtctgaca gaaaatgat atatgacgag aagatagcag 840
agttcaaaag acaacttgca gaggaggaa ccatggatac agatcaaggt aatagtattg 900
taagtgtcat tcaagcagaa gttgccaaaa atcagatgct tattgaaga gaagtacaga 960
aattaaaaag atacaagatt gagaatatca gaaggaagca taattattct cctttcaata 1020
tggaaattgtt aaagacttta aagctcaggg aaaccaaagt accactagta gaaaaggcaa 1080
aagaaaaaca gaaccaaaag aaagctcagg aaaccaaagt accactagta gaaatagta 1140
cacattttct cttctgcaca tattttcatg gaacacatta tgtataaaga acttagagca 1200
acatcctaag ttgctcagtg cactgtttggc aatagtggca gctgtcttgc tctttaatgc 1260
atggtattcat aaactttctt cctacctgca tcatgtgcat gtatgtcata ttaaatgaaa 1320
gtgatatttaa gatgtcttcc ccaaattcca ttatttgaca ttgagtctga caactgttag 1380
ttttctgggt gtccaaactac catatgaagc tagaaaaagc acaaacgata ttccattatc 1440
gtaattttaa tacttaaaat ttgcaattgt cagatcttga ttaacctgtg ttactatttt 1500
cttctcatca ttaacggaaa aaaaatcagt atttctatct ttgatatcta agtgttttga 1560
ggatttttaa actgaatttt atctgctata ccagttattt gagaaagtat gattttcaag 1620
taaatcattt aaaaaggaca aaagtataat ttccagtgat tttcactgct gtcagtagaa 1680
aagtaataaa catctcaatt ttaaaaaaaa aaaaaaaaaa aaaaaaaa 1728

```

<210> 395

<211> 1348

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006117

<400> 395

```

caggtggcgt acttggtctg gagactggcg cggcgcttct gtcagagtc tctgcaggtc 60
actagtttcc cggtagttca gctgcacatg aatagaacag caatgagagc cagtcagaag 120
gaactttgaa atccaatgaa tcaagtgaag cctctgaaaa aggatccagg aacaggaagt 180
aagctaaaaa tctacgcgct atataagcac gccactgaag gaacctgtga catgcccaaa 240
ccaggtgtat ttgactgat caacaaggcg aatggggagc catggaatgc ccttgaggag 300
ctgcccaagg aagctgccag gcagaaactat gtggatttgg tgtccagttt gagtccctta 360
ttggaatcct ctagtccagg ggagcctgga acagacagga aatcaactgg gtttgaaact 420
ctgtgtgtga cctccgaaga tggcatcaca agatcatgct tcaaccggcc caaaaagaaa 480
aatgccataa acactgagat gtatcatgaa attatgcgtg cacttaaaag tgcacgaaga 540
gatgactcaa tcatcactgt tttaacagga aatggtgact attacagtga tgggaatgat 600
ctgactaaat tcaactgatat tccocctggg ggagttagag agaaagctaa aataatgccc 660
gttttaotga gggaatttgg gggctgtttt atagattttc ctaagcctct gatgtgcagt 720
gtcaatggtc cagctgtggg catctccgct accctccttg ggctatttca tgcctgtlat 780
tgccttgaca gggcaacatt tcatacacca tttagtcaac taggcocaaag tccggaagga 840
ttgctcctett acacttttcc gaagataatg agcccaggca aggcaacaga gatgtctatt 900
tttggaaaaga agttaacagc ggggagaggca tgtgtccaag gacttgttac tgaagttttc 960
ccttagatgca cttttcagaa agaagctctgg accaggctga aggcatttgc aaagcttccc 1020
ccaaatgctt tggagatttc aaaaagagga atcaggaaaa gagagagaga gaagactcac 1080
gctgttaatg ctgaagaatg caatgtcctt cagggaagat ggctatcaga tgaatgcaca 1140
aatgctgtgg tgaacttctt atccagaaaa gatgaccact gatgaccact aactcaatca 1200
aaagcatgct caaggaaagga tgtgctgtta cctctgattt ccagtactgg aactcaatca 1260
gcttcattgt gcttttga taagttagaa atcaattaca atgatgata ttcactacag 1320
ctctgagtaa taaaagttt tgtaaaaa aaaaaaaaaa aaaaaaaa 1348

```

<210> 396
 <211> 3106
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AK000745

<400> 396
 cggcgccggc ccccgacgac ctggccttgc tgaagaatct cgggagcgag gaacagaaga 60
 agaagaaccg gaagaaactg tcgcagaagc ocaaaccaaa tggggcgagc gttgaagtgg 120
 ctgagggtga agctgttoga acacctcaaa gtgtaacagc aaagcagcca ccagagattg 180
 acaagaaaaa tgaagaagtc aagaaaaata agaaagaatc aaagtcagat gctaaagcag 240
 tgcaaaacag ttccagccat gatggaaagg aagttgatga aggagcctgg gaaactaaaa 300
 ttagtccacg agagaaacga cagcagcgta aacgtgataa gtgtctgact gattctggtt 360
 cattggatgc aactatccct gggatagaaa ataccatcac agttaccacc gagcaactta 420
 caacgcgctc atttctgttt ggttccaaga agaataaagg tgattctcat ctaaatgttc 480
 aagttagcaa ctttaaatct ggaaaaggag attctacact tcaggtttct tcaggattga 540
 atgaaaaact cactgtccat ggaggaggct ggaatgaaaa gtctgtaaaa ctctctccac 600
 agatcagtcg aggtggagg aagtggaact ccgtttcacc tgctctgca ggaagaggga 660
 aaactgagcc atctgcctgg agtcaagaca ctggagatgc taatacaaat ggaagagact 720

ggggaaaggag ttggagtgc cgttcaaat tttctggcat tgggtctact gctgagccag 780
 tttctcagtc taccacttct gattatcagt gggagtgtag ccgtaatcaa ccctatatcg 840
 atgatgaaat gtctgggtta aatgggtctgt ctctctgtga tcccaactct gattggaattg 900
 caccagcaga agagtggggc aattgggttag acgaagaagg agcttcaact ctaaagtccc 960
 aggaaccaat tctgtgatg caaaagggtct cagatgatga taaagaaaaa ggagaggagg 1020
 ctctctcaac ttgggaatcc aaaaagaaaa gaagaagcaa ggtgaagata 1080
 actctactgc acaggacaca gaagaattag aaaaagagat tagagaagac ctccagctga 1140
 atacctctaa aaccggtcca aaacaggaaa agcttttttc attgaagacc atagcactta 1200
 gtgatccagc cgaagtactc gtcaaaaaaa gccagcctat caagactctt ccaactctga 1260
 ctctccacga gccactctga atcttatcaa aaagtgtatt tgacaagagc tcttcccaag 1320
 tgccgccaat actacaagag acagataaat ccaagtcaaa taccagcaaa aatagtgtgc 1380
 ctctctcaca gaccaagtct gaaactagct ggggaatctcc caaacaataa aaaaagaaga 1440
 aaaaagccag acgagaagcg tgaaattttt tttctcgaat tggacatgtg ttgcaaaaca 1500
 ctgtctctga agattatgct gtttatgcaa taatttgtga acatgtacag agttttatat 1560
 aaatttaaac caatttttaa acaaaaactg cggacacacc agaataact ataatatact 1620
 aaagttaatt tatgaaatta agaggtcagc agaatatact cagtgtatga agacacttgc 1680
 gaaagtcttt ttaatagaac aagaacgcat ttaatttaag aatattatcc tggtttaaca 1740
 acagtgcctc gtttacaaca gattgtgcc cttctcatct tatctcatct cacttgagct 1800
 tctgtattaga aaggaggttg cctacagatt agtaagcaat tccctggact ttaatgcacg 1860
 aactgtgacc atttgaatct gttttatgct taaatcaaa tgctttgac aaatgcataa 1920
 ccctgcacat ttgtacatat ttgttggtag caattttgat taaagaaatc acaagtgcac 1980
 ataaaaagct atttatcatt tgtttaacta aactgtcatg gtttagttta caatttttaa 2040
 aaagtcttta aaactactga aatgcagttg atggctttat ttggctttat aagttatttt 2100
 tgatagtcct acattacttg aattgttcaa agtacaagat atttttaatt aagaaaagtg 2160
 aactatcata atttgtttta tacattttaa ctatagactc ataaataatg ctattgttta 2220
 tgatttgtaaa atcttcaggc aaaaatcaat ttacattttt ccttccctca gcaattactt 2280
 ttttccagct tcaactcttc ttatgttact gactttaaaa atgaatcatc 2340
 tcacaaactt tctgtatatg atggagaatg aaaaactaga gtcagacagc ttaattatga 2400
 attgtcaaca cctccagtta tcagggaatac tgctctaacg tgctagtgcg 2460
 agaattatga tcaatttctt gaaggagatt catgttttta taagaatttt catgttaata 2520
 ttgcaattgt ggtcaaaata gaaacgtttc ctgcttgaaa ttatatgat ttaaatgatg 2580
 tgtgagatgt ttcacattt tcaggcactg tgtaattcta ttgtaataaa ctggcaggta 2640
 tctttgtaac tataaatagt gcatgtcac ccatgtacac tgtaaatagc cttaccaca 2700
 cgtgtgtgac aaggaccata attaacatca cttagtgaat tgtgataag aaaaaaagc 2760
 catgatttat tcatgttgat tggctgtttt ttatgtggcg ccaagaacga actgatttaa 2820
 cagctgtaac caatgtgtact gatctatcca tccaatgttg tctatattat tgaactgtgt 2880
 tcaacagtat tgcgtgtgca gactaggaaa gctaaacgaa caaattggtt ttgattttgc 2940
 tgaagactgg cttatttaac ggaacagctt gctatttaac gattattaac ttttatcagg 3000
 tgtaaacatc tgtttcagga acatggcagt atgtttacat gtcagaagtt ttgttttaatt 3060

ctatggtatt tctaaattga cttgttttaa taaattcagc aaatgg 3106

<210> 397
<211> 556
<212> DNA
<213> Homo sapiens

<300>
<308> Contig40831_RC

<400> 397
ctgtttttat catgtaattt tcagaaaaata tgtgatccat aaagattaaa agaaagtgtg 60
attaagtctg cgagcttttag tattaacttg aaataaaaaa tggcaagctt tccacgtcct 120
cctttatttc cacaatccat atgtacgagc tagattccag tcagaaacttc cacaataact 180
tcaactcttg gtacgagcgg ttataaatta cgctttgct aatttgcgtt gtcccaaac 240
aggagaaaca ttaccacaaa aaaagtcagt ttcactctgc agtgttcccg cagcaacat 300
attagctcgg aagaataaaag ctcccttgta gtagtgcgt ctattgaatt acaatgtaaa 360
tgttgatgcc tccctggaaa acgtcaaatg acaacaggtt tacagaacga attcagtggt 420
tgtttgcagc tctggaaatcc caggctgcag agttgagatc gaagaggtgc gtctggccat 480
cttagcgttc attcaaccaa attcctctac tccatcaaaa ccacgccact cccacactgc 540
ccagctcgta aaagga 556

<210> 398
<211> 893
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003239

<400> 398
gagttttagc agggagggat atgcatgtac acatatgcat atacatacag gttttatttaa 60
gtgattggat ctacactttt ggaggtctatt tgaagtatct tatcaaaaata aaacgcattt 120
taagcaagag gttaaaaaag cttttcagac ccgaaggcca tttaatacag taaatcttac 180
ttgggtagtt tagcaaacat tttttaaaaa ccacatccaa cagattggtt agtatttagat 240
acacagattt ggtttcaaaag ctgaaatacc cctggtggaa taaatcttct catgaaattc 300
agtatatatta tgccacaaaa cagggacact tatctattga caacttaaat attaaactac 360
taggctacct ggtgatataa ataggaaaaa aacagtgggtc atggattagg agtgagaaga 420
ctatgacacc agctacactt ggactgggac cgcaaggagt gttaaaatct atcctagaat 480
ttaaaacagag aaaaaatgca aatattccac tcccattcg gcccaaacca tcccccaagg 540
aatacacccc caccacccc ccacatgagg tttactaacc agctgctaca gatgttacta 600
tgtcactaat taacaatgat cacagaaaat ttcctatacc gtgtcatacc agagtttttt ttatcataca 660
caatgcatta aaagtctcat ccgctggttc caaaaagcta ccgtcagaag gcaaaactgac 720
cattgagctc tctctggcat gctcccaact cagctaaagag ccatgaggca ttcatttgta 780
cgtcccagcc gggcggtgca gctcctact cctggggctc ttgctgggtca act 840
893

<210> 399
<211> 2470
<212> DNA
<213> Homo sapiens

<300>
<308> NM_014791

<400> 399
ttggcgggag gaagcggcca caaccgggag atcgaaaaga ttcttaggaa cgcgtacca 60
gcgcggtctc tcagacagc aggcctctgt cctctgtcgc ggcgcgctc agcgtgcc 120
tcgcgccttc aggttctttt tctaattcca aataaactgt caagaggact atgaaagatt 180
atgatgaact tctcaaatat tatgaattac atgaactat tgggacaggt ggctttgcaa 240
aggtcaaatc tgcctgccat atccttactg gagagatggt agctataaaa atcatggata 300

aaaacacact	agggagtgat	ttgccccgga	tcaaaacgga	gattgaggcc	ttgaagaacc	360
tgagacatca	gcataatgt	caactctacc	atgtgtctaga	gacagccaac	aaaatatcca	420
tgggtctctga	gtactgccct	ggagagagagc	tgtttgacta	tataatttcc	caggatcgcc	480
tgtcagaaga	ggagaccggg	gtgtctctcc	tcagatagat	atctgctgtt	gcttatgtgc	540
acagccaggg	ctatgctcac	aggacacctca	agccagaaaa	tttgcgtgtt	gatgaatcac	600
ataaattaaa	gtcgtgtgac	tttggctctc	caagggtaac	caagggtaac	aaggatcaac	660
atctacagac	atgctgtggg	agtctggctt	atgcagcacc	tgagttaata	caaggcaaat	720
catatcttgg	atcagaggca	gatgttttga	gcactggcat	actgttatat	gtctcttatgt	780
gtggatttct	accatttctg	gatgataatg	taatggcttt	atacaagaag	attatgagag	840
gaaaatagga	tgtttcccaag	tggctctctc	ccagttagcat	tctgtctctt	caacaaatgc	900
tgcaagtgga	cccaaaagaaa	cggtatttcta	tgaaaaatct	attgaaccat	ccctggatca	960
tgcaagatta	caactatcct	gttgagtgcc	aaagcaagaa	tccttttatt	cacctcgatg	1020
atgattgcgt	aacagaactt	tctgtacatc	acagaaacaa	caggcaacaa	atggaggagt	1080
taatttcaat	gtggcagtat	gatcacctca	cggtcaccta	tcttctgctt	ctagccaaga	1140
aggctcgggg	aaaaccaggt	cggttaaggc	tttcttcttt	ctcctgtgga	caagccagtg	1200
ctaccaccatt	cacgacatc	aagtcaaaata	attggagtct	ggaagatgtg	acgcgaagtg	1260
ataaaataat	tgtgcgggga	ttaatagact	atgattgggt	tgaagatgat	ttatcaacag	1320
gtgctgctac	tcoccgaaaca	tcacagttta	ccagtgactg	gacagaatca	aatggcggtg	1380
aatctaaatc	attaactcca	gccttatgtca	gaacacctgc	aaataaatat	aagaacaaag	1440
aaaatgtata	tactctaaag	tctgctgtaa	agaatgaaga	gtactttatg	tttctgagc	1500
caaaagactcc	agtttaataag	aaccagcata	agagagaaat	actcactacg	ccaaatcggt	1560
acactacacc	ctcaaaagct	agaaaccagt	gcctgaaaga	aactccaatt	aaaataccag	1620
taaattcaac	aggaacagac	aagttaatga	caggtgtcat	tagccctgag	aggcggtgoc	1680
gctcagtgga	attggatctc	aaccaagcac	atatggagga	gactccaaaa	agaaaggagag	1740
ccaaagtgtt	tgggagcctt	gaaaaggggt	tggataagtg	tactactgtg	ctcacaggga	1800
gcaaaaggaa	gggttctgcc	agagacgggc	ccagaagact	aaagcttcac	tataatgtga	1860
ctacaactag	attagtgat	ccagatcaac	tgttgaatga	ataaatgtct	attcttccaa	1920
agaagcatgt	tgaacttgta	caaaagggtt	atacactgaa	gtgtcaaaac	cagtcagatt	1980
ttgggaaagt	gacaaatgcaa	tttgaattag	aagttgtgcca	gcttcaaaaa	cccgatgttg	2040
tgggtatcac	gaggcagcgg	cttaaggcgc	atgcctgggt	ttcaaaaaag	ttagtgggaag	2100
acatcctatc	tactctcaag	gtataatgta	tggattcttc	catcctgcgc	gatgagtggt	2160
gggtgtatcc	agctctacata	aagactgtta	tgatctgctt	gatttttaag	ttcatgtgga	2220
ctaccaactt	agctatctta	agctatctta	agaccaaat	ctcttttgtt	ttcaaaaaaa	2280
gatattattt	tgtgttgtaa	tctaaatcaa	gcccatctgt	cattatgtta	ctgtcttttt	2340
taatcatgtg	gttttgtata	ttaataattg	ttagactttc	tagattcaact	tcocatgtgt	2400
aatgtaagct	cttaactatg	tctctttgta	atgtgtaatt	tctttctgaa	ataaaaccat	2460
ttgtgaatat						2470

<210> 400

<211> 2212

<212> DNA

<213> Homo

<300>

<308> X05610

<400> 400

ggggaacgag	gccacactgg	gagccacgga	cttcaggggg	tcccaggcat	cacacccccct	60
tccaacatct	ctgggggcac	tggtgacaaa	ggggcgccag	ggatatgttg	cctgaaaggt	120
tatccgggcc	caccagggcc	accaggttct	gctgctcttc	ctggaagcaa	aggtgacaca	180
gggaacccag	gagctccagg	aacccacagg	accaaaggat	ggggccggga	ctccggggccc	240
cagggcaggc	ctgggtgtgt	tggtctccca	ggagaaaaag	ggccaggggg	tgaacagacc	300
ttcatgggga	acactggacc	caccggggcg	gtggcgacga	gagggcccaa	gggacccaag	360
ggagaccagc	gattccctgg	tgcccccg	actgtgggag	cccccggaat	tgcagggaatc	420
ccccagaaga	ttgccatcca	accagggaac	gtgggtcccc	aggggaggcg	aggccccctt	480
ggggcaccgg	gggagatcgg	gccccagggc	ccccccggag	aaccaggttt	tcgtggggct	540
ccagggaagc	ctggggccca	aggaagaggt	gggtgtctct	ctgttcccg	cttccgggga	600
gatgaaggac	ccataggcca	ccaggggccg	attggccaa	aaggtgcacc	aggccgtcca	660
gggagccggg	gcctgcggcg	tatgccaggc	cgcagcgtca	gcactcgcta	ccctcgtgtg	720
aagcacagcc	agacggacca	ggagcccatg	tgcccggtgg	gcataacaaa	actctggagt	780

```

ggatacagcc  tgctgtactt  cgaggggcag  gagaaggcgc  acaaccagga  cctggggctg  840
gcgggctcct  gcttggcgcg  gttcagcacc  atgccttcc  tgtactgcaa  ccttgggtat  900
gtctgtactat  atgcagcccg  gaacgacaag  tctacttgcc  tctctaccac  tgccgctgtg  960
cccattgatcg  ccgtggccga  ggacgagatc  aagccctaca  tcagccgctg  tcttgtgtgt  1020
gaggcccccgc  ccatcgccat  cgcggtccac  agtcaggatg  tctccatccc  acactgccca  1080
gctggytgagg  ggagttttgt  gatcggatat  tctctctcca  tgccacacgc  ggcgggagac  1140
gaaggcggtg  gccaatcact  ggtgtccacg  ggcagctgtc  actacgcca  ccgcgccaca  1200
ccattcatcg  aatgcatagg  aggcgcggcg  acctgccact  actacgcca  caagtacagc  1260
ttctgtgtga  ccacacttcc  cgagcagagc  ttccagggtc  cgccctccgc  cgacacgtgc  1320
aaggccggccg  tcatccgcac  acacatcagc  cgctgcgaag  tgtgcatgaa  gaaacttcta  1380
gcggcgcgct  gccaggaagg  gccattttgg  tgccttattc  taacttatta  cctcagggtg  1440
caacccaaaa  ttggttttat  ttttttctta  aaaaaaaaa  aaagtctaac  aaaggaattt  1500
gcattccagca  gcagcactta  gacctgccag  ccaactgtac  cgagcgggtg  caagcactgc  1560
gggtccctgag  aggcgaagccc  tgcccacaga  aagccaggag  cagccctggc  cccactcagc  1620
cctgctacga  cgcaccgcct  gaaggcacag  ctaaccactt  cgcacacacc  cactgtaacca  1680
ctgcactttc  caatgccaca  gacaactcac  attgttcaac  tcttctcgg  ggtgggacag  1740
acgagacaac  agccacacag  cagccagccg  tgcccgaggg  ctcgaggggc  gtaggggttc  1800
agggcaccgt  ccccaacaga  gggcccctg  ggtggcctg  ccttctcttc  tacgccaatg  1860
ttatgctacg  tccatgttct  cccaataacc  gttgatgtga  attattttta  aggcacaaat  1920
gtgctctttaa  ttttaaaaaa  cactgataat  cacactcgg  taggtcatt  ttttgcacca  1980
tccctatcga  ccaactgggt  tggcaaaact  caggcagaag  tggagacct  tctagacatc  2040
attgtcagcc  ttgctacttg  aaggtacacc  ccatagggtc  ggaggtgtcg  tccccactgc  2100
cccactctgt  cctcgagatt  taacccctcc  actgctggg  gtgagctgta  ctcttctgac  2160
tgccccctcc  tgtgtaacga  ctacaaaata  aaactgtgtt  ctgaattatt  tt  2212

```

<210> 401
<211> 4221
<212> DNA
<213> Homo sapiens

<300>
<308> NM_016448

```

<400> 401
cgataacgat  ttgtgttgtg  agaggcgcaa  gctgcgattt  ctgctgaact  tggaggcatt  60
tctaagcatt  ttctctcagc  tgaggctttt  cctccgaccc  tgatgctctt  caattcgggtg  120
ctccgcacga  cccagcttgg  cgtctctgaga  aatgatgggt  ctccacaata  ccctcttcaa  180
tccctctctg  ttgcttatca  gtgcagtggt  aatgatgaac  acacttctta  tgggaaacaa  240
ggagctccag  tctctctctt  tggatgtacc  ttctctctcg  ctcccaatat  ggaactatga  300
ctagcagttg  ccaatgaaga  aggtctttgt  cgattgtata  acacagaact  acaaagtttc  360
agaagaaggt  gcttcaaaaga  atggatggct  cactggaatg  ccgtctttga  cctggcctg  420
gttctcgttg  aacttaaaact  tgttacagca  gcaagtgatc  aaacagccaa  attttggagc  480
gtaaaagctg  gtgaagctgat  tggacatgc  aaagtgctac  tgtacgggtg  gaagagatgg  540
gccttttcta  agttttagaa  agctgtatct  ggtgtttata  ggcaagtga  caactattgt  600
gtctgggata  ccaggtgcaa  caaaaaagat  ccttcaaaac  ccaagaagaa  tcaaatcagt  660
ggagctcaca  atacctcaga  caagcaaaac  ccttcaaaac  ccaagaagaa  caggaattca  720
aaagacttg  ctctctctgt  ggaattccag  caaagtgtta  ctgtggtcct  ctgttcaagc  780
gagaatacct  tagtctcagc  aggagctgtg  gatgggataa  tcaaatgatg  ggatttacgt  840
aagaattata  ctgcttatcg  acaagaaccc  atagcatcca  agtcttctct  gtaacctaac  900
agcagcactc  gaacaaactg  atattcaagt  ctgattttgg  atctccactg  ctactcttta  960
tttgctaatt  gactcagaca  taacatctac  atgtttaata  tgactgggtt  gaagactctt  1020
ccagtggtga  ttttcaatgg  acaccagaac  tctacctttt  atgtaaaatc  cagccttagt  1080
ccagtgacc  agtttttagt  cagtggctca  agtgatgaag  ctgcctacat  atggaagtc  1140
tccacaccc  ggcaacctcc  taactgtctc  ctgggtcatt  ctcaagaggt  cagctctgtg  1200
tgctgggtgc  catctgactt  cacaaagatt  gctacctgtt  ctgatgacaa  tacactaaaa  1260
atctggcgct  tgaatagagg  cttagaggag  ataccaggag  gtgataaact  ttcccaggtg  1320
ggttggggct  ctcaagaaga  aaaagatgca  agacctggcc  tagtaaacgt  aacgagtagc  1380
cagagtactc  ctgccaaaag  ccccagggtta  aagtgcatac  catccaatc  tactccatgc  1440
tcgcagactt  gtgcccacag  ctgtgctgga  gacctccctc  tctcttcaaa  tactctcagc  1500
ttctctatta  aaactctccc  tgccaaaggc  cggctctcca  tcaacagaag  ctaactctgt  1560
tctctcgtct  ctcccagacc  accttactct  ttcaagatgt  cgattagaaa  ctgggtgacc  1620

```

cgaacacctt cctcatcacc acccatcact ccacctgctt cggagaccaa gatcatgtct 1680
 ccgagaaaag cccttatctc tgtgagccag aagctatccc aagcagagcg ttgctctgag 1740
 tctagaaata gagtaaaagag gaggctagac tcaagctgtc tggagagtgt gaaacaaaag 1800
 tgtgtgaaga gttgttaactg tgtgactgag ctgtgagcgc aagttgaaaa tctctatttg 1860
 gatctgtgct gccttctgttg taaccaggaa gaccttagta aggaactctct aggtcctacc 1920
 aatatcaaga aaattgaagg agctggatcc agtatctcag agcctccgtc tccatcatgt 1980
 ccgtatgctt cagaaaagctg tggaaacgcta cctcttccct tgagacctgt tggagaaggg 2040
 tctgaaatgg taggcaaaaga gaatagtctc ccagagaata aaaactggtt gttggccatg 2100
 gccagcaaac ggaagcgtga gaatccatct ccacgaagtc cgtcatccca gacaccaact 2160
 tccaggagac agagcggaaa gacattgcca agcccggtca ccatcacgcc cagctcattg 2220
 aggaaaatct gcacatactt ccatagaaag tcccaggagg acttctgtgt tccctgaacac 2280
 tcaacagaat tatagattct aatctgagtg agttactgag ctttgggtcca ctaaaaaag 2340
 ctgagctttg gtccactaaa acaagatgaa aaatacaaga gtgactctat aactctggtc 2400
 tttcaagaaag ctgccttttc atttttagac aaatctttt caacgctgaa atgtacctaa 2460
 tctggttcta acccataaat gtatatgcag gtcccagag atgaatgctg ttttaaat 2520
 tcatataagt aatttgtcac tctagcattt tgaatgaata gtcttcactt tttaaattat 2580
 tcatctcttc tataataatg acatcccagt tcatggaggc aaaaaaag ttctttgtta 2640
 tccctgaaact tctatgtctc agtggaagtg atctgccagc cacagcatga ggccctgtgaa 2700
 ggctgactga gaaatcctct gctgaagacc cctggttctg ttctgcctcc aacatgtata 2760
 attttatttg aaatacataa tcttttctac atgtctttgt ggggtttttt ttaagtatgt 2820
 gtaaaaatgt gatgctcaga taagtacatt tatatcagtt cagtgtaaaa atgcagctc 2880
 ttgagttaaa gtacatttta ttttaaatgc agtgataaat gtcaactctt cggagaact 2940
 aggagaacaa accagaaaag ctgtgtttgt ctttttctc tcaaatatat ccccgtagt 3000
 agatttcaag tcccatggtt ttcaccaagc aatctgctat gtcagccaac ccaactcac 3060
 tttctacagg agcttatgat ttttgccatt tactagagga agatgtttta tgaacctaat 3120
 ttggggtttg aattcaggtg cagtcatcag tcttttaggg ctgcgaatgt tttaaaaaaa 3180
 ataatgtcac agattttaa gaaaaagtga tgattttcta ttgatattt tgtacagaa 3240
 tatagctctt aactgaaaa ccaagaaccg aaacataaat cttgagtttc ttttcatgta 3300
 cataaaaagc aatagccttt tagtatagat agccctgagc caaaaagtaa tagaatttt 3360
 tctagatatt taataccag agtgatataga ctgactctaa ctgtaaatag tgaataatg 3420
 cttcaaacat cctcccttta tccaacaatt atgtatcagt gatcttgaac cattgtttta 3480
 tatttttcac ctttgaacc tcatggaaag aggctttaca taacttctat gtactattta 3540
 cttagaaggg agcccccctc cagtcatgaa acttcaattg tttatccat atccctgagg 3600
 actgtgtaga ttttatgtca gttctgtgta cagtttttgt cagtttttgt cattatttga 3660
 aaatctatc tgacaaactt ttaattcctt tgactctata agttaaagct gtaacaaact 3720
 aaattgcatg taatacagta gcatagtttt atccaggag atccaggaat ggaagccata 3780
 gaattgctct ggtcaaaacc aagcacacca tagccttaac tgaatatata ggaactctgc 3840
 ctaactgtct taatttttgt gtttgttttt tgactgttgg gctttgggaa ggtgttattt 3900
 atgaccaata tctgccagta acgctgttta tctcacttgc tttgaaagcc aatgggggaa 3960
 aaaaatccat gaaaaaaaag agattgata agtagatgat ttgttttga tccctacca 4020
 tctctctggca gccctactga gtgaaattgg gatacaattt gctgtcagaa attataccga 4080
 gtctactggg tataactagt ctcacttgga aagctatgac agacttttga gtgcacagg 4140
 tcaactgtaa tgagataatt atccctgcct gtgtccatgt cagactttga gctgatctct 4200
 aataataaag cttttacct t 4221

<210> 402

<211> 3224

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018401

<400> 402

gtccacatc	ccgcatccgg	catcccagcg	gccgggcatg	tagcagcgcg	agcaacggcg	60
gaatatggcg	gggaaccact	cccacaagcc	ccccgtgttt	gacgagaatg	aggaagtcaa	120
ctttgacctt	ttccagatct	ttcgggccat	tggttaaagg	agttttggaa	aggtatgat	180
cgtgcagaag	cgagacacta	agaaaaatga	tgcaatgaag	tacatgaaca	agcagaagtg	240
catcagagag	gatgaggttc	ggaatgtttt	ccgggagctg	cagatcatgc	aaggcgctga	300
gcaccccttc	gtgtccaact	tggtgtaact	cttccaggat	gaggaggaca	tgttcaatgt	360
ggtggacctg	ctcctgggag	gcgacctgcg	ctaccatctg	cagcagaatg	tgactttcac	420


```

agaggggact gtgaaactct acatctgtga gctggcactg gccctggagt atcttcagag 480
gtaccacatc atccacagag acatcaagcc agacaatata ctgctggatg aacacggaca 540
tgttccacat acagacttca acatagcgac ggtagtgaag ggagcagaaa gggcttccct 600
catggtctgg ccacagccct acatggctcc agaagtatc caggtgtaca tggacagagg 660
ccccggatac tctgacctcg tcgactgggtg gtccctgggc atcacagcct atgagctgct 720
gctggctggg ccggcgctag aaatccactc ggtcacgccc atcgatgaaa tctccaacat 780
gttcaagggt gagcgtgtcc actactcctc cactctgtgc aaggggatgg tggccctgct 840
gaggaagctc ctgaccaagg atcctgagag ccgctgtgtc agccttcatg acatacagag 900
cgtgccctac ttggccgaca tgaactggga cgcggtgttc aagaaggcac tgatgcccg 960
ctttgtgcc aataaaggga gttgtaactc cgaatccaca tttgagcttg aagagatgat 1020
tctagaatcc aagccaactc acaaaaggaa gaagcgattg gcaaaagaaca gatccaggga 1080
gggcacaaag gacagctgcc cgtgtaatgg acacctgcag cactgtttgg agactgtccg 1140
ggaggaattc atcatattca acagagagaa gctcaggagg cagcaggagg agggcagcca 1200
gctcttggac accgacagcc gagggggagg agcaggccca agcaagctcc agggcgggtg 1260
caacaacaac ctctcaccc acacctgcac ccgtggctgc cgcagctgag cccactctgt 1320
ttgctgtcca acaggactgc actcgtctct gccctgccca cccagagccc ctctttgtgc 1380
ctcgtatggtc cctgtctcac cctgaaaac atcagatgac catgactgat tcactgtga 1440
gctgggaagc ctgggttctg gtcccatctc catcgtctgc tcactgtga cctcagaca 1500
gtcacgcccc ctctgtgccct ccgttttctg catctgccaa aggggtttaa cactctcgcc 1560
ccaactcaaa ttacaagatt atggggagaa cccaattagg taggaaacat gaaaaacctt 1620
tgatatttat aaaaatcttt ttacgtgcaa atataacct taattattga agtgaccccc 1680
atctcccaaa gcaatcaaac cgtcatgact ttgcaatttg gcaatcctta gctttagata 1740
gggcaacttc agaaaacaca gccctgacag caaaaataag gtctgatatg gtgccccct 1800
ctgggaaaac aacgtgccca aatcctggag caaaacctga agtgctctca tgtgcatct 1860
ctgacagccc acagctctga gcttgaaga tggctgcagca tgcagaccag acttgcccc 1920
aaggtctcag cgtgcggctc tcaactctcc cctcatttaa gaagactatc cttaactttt 1980
agtttcagca gtctccacca ccaccatctc ccagtgctg ggatggcaca caggtgtcca 2040
tccagatgag agttgggtcg ctgagcatg gctactctct cagagtgtaa caggaccccc 2100
atccaactgg ccggaagccc cagacctgca gcagaaactct ccaactctct atcagctttc 2160
aggggtttct ctctgggaaa ggggtgaaaa agctgtgttc agatctctct tacagagagt 2220
atccaatcgg tattgttggg ggggtccctc attatacaa taggaagcat ggggtgctag 2280
aaaagtttat tcaggaggaa aatgggttca cacaaaaagc aaactacat ctgctctgt 2340
cagggagaag cttgcccttg aactggaaga tgttgggatg agcagggaaa gcttagactt 2400
tggagtcaag tttgtgttca gaatccagcc ctgctgctca ctaactaact cgtatcgtc 2460
aggcaaaagc tgcaatcgct ctgaatggca gtttctcat ttttaaacag ggataataaa 2520
actaatattg caggggagtt acagggttaa ataatctct gtgtgtaacc ccaagcattg 2580
gatgactact agaattggct tttttgtcag cataatctgc atcattattt agatactttc 2640
ttccttcaact accccagcag gtcagtttct tgtgcaaaaca aacctgttta ggaactcttc 2700
aaatgttctt cgtggggtct ttgatatttg tttgttacat cctgctgaag ttcgactgtg 2760
tttttttttt tctcaacaac ttcoattttt cactttttac atgtactaact aatccttggg 2820
gctgtccatg tcatctctta gatttcttaa aagacatttt aatgtatggt taggtttttt 2880
attttttatt tttaaaaag aaatagtca gttttcctc ctttcaacog agactatttc 2940
tggattgtgt gctcctcgct agttgaactg ttttgcaaac ttttcttacc ttcactgtcc 3000
catcaacaac cgtcctgctc cccacctccc ccaggaataa aggggctgct tctctcctct 3060
actgtgaccc tggaggctct taagatgatg atgtgttttt ttatttggct gagttcaacga 3120
atttagggga gggagctgaa gtgcgccatg gaacaccaga tttctgggtt ctgttcaagt 3180
tggcatttct tgtttggaat aaactatttc ttggacattc cttc

```

<210> 403

<211> 2460

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000788

<400> 403

```

aaagtcaaac cccgacaccg cggcgggccg gtgagctcac tagctgaccc ggcagggtcag 60
gatctggctt agcggcgccg cgagctccag tgcgcgcacc cgtggcgccc tcccagccct 120
ctttgcggga cagactctgg gccgcaccaa gactaaggaa tggccacccc gccacaaga 180
agctgccegt ctttctcagc cagctctgag gggaccacca tcaagaaat ctccatcgaa 240

```

```

gggaacatcg ctgcagggaa gtcaacattt gtgaatatcc ttaacaacatt gtgtgaagat 300
tgggaagtgg ttcttgaacc ttgtgccaga tgggtcaattg ttcaaaagtag tcaaaagtga 360
tttggaggaac ttacaatgtc tcagaaaaaat ggtgggaattg ttcttcagat gatgtatgag 420
aaactcgaaac cactgtcttt tacctcccaa acatatgcct gtctcagtcg aataaagact 480
cagcttgccct ctctgaatgg caagctcaaa gatgcagaga aacctgtatt attttttgaa 540
cgatctgtgt atagtgcacg gtatattttt gcatctaatt tgtatgaatc tgaatgcattg 600
aatgagacag atgtggacaat ttatcaagac tggcatgaact ggatgaataa ccaatttggc 660
caaaagccttg aattggatgg aatcatttat ctccaagcca ctccagagac atgcttacct 720
agaattatatt tacgggggaa aatgaagag caaggcattc ctcttgaata tttagagaag 780
cttcattata aacatgaaag ctggctcctg cataggacac tgaaaaacca tctcgattat 840
cttcaagagg tgcctatctt aacactggat gttaatgaag actttaaaga caaatatgaa 900
agctcgtgtt gaaagggtcaa agagtttttg agtactttgt gatcttgcgt aagactacag 960
gcagccaaat ggttcagat acttcaagctt tgtgtatctt cgtaaactta tattaatata 1020
agtttcttta gaaaaaccaa gtttttaact gttttttgtt taaggaaaaa agatttttaa 1080
aatgaactct atgcaaaaact ttttgatcag ttctttttt ttgtttttt ttttaaaaaa 1140
gacattctaa gacaagacac ttatttctca tagcaggaaa tgtagaggta gatggttcca 1200
gatcagcat agtgactaaa ctacattata aaagatccag ctctctctg tcatccctc 1260
cttttgcct cctcagcagg ttggcttttt tccttgggtg ctctcacttc gttgggtgac 1320
agtttcttaa acttgaaagct ttaagtgtac atagtaaatg ttagtgtgtc ctggtgaagt 1380
tagtgcacct attaaaagtt gcaaaagtga attaaaggaa tccctagaat aaggattctg 1440
aagttttatt ttaaaattat atcttcttaa cagtttagtc ccactctta ctctcgtcct 1500
cagtcgtcct tctctactgt ctggattaat taggcagcct gctataaagt taaagtcaca 1560
cattcttatt ttgcaaacac tgtgattact ctgtgtttg tagttttagt tgcctttagt 1620
ggttctgctt ttaagttttt ctctttttca gacaaattac tgataaaaaa gatattgctc 1680
tatattgaa atactctgaa agcattattt tttgttgaa aggaataaaa ataatgaag 1740
acagaggcta gaaagatccc attaattaat gagacacac taactacta tctctaaacc 1800
atctatgtga atatttgtaa aataatgaa tggactcacc ttagtctgt atataaatat 1860
attttctttc tagtttgttt agttaagggtg tgcagtgatt ttcctgtga ttaaaccttt 1920
ccattttacg tttttagaaa ttttatgtat ttttaataaa gggggaagag cattttcaac 1980
tttaaacact tatttttctt tccaagtcac ttttgtttt ggtttcttat caaagatga 2040
taatttagtg gattaaaccg tccagcagca ctgactcttg caaaggagac ttaatttcaa 2100
atctgtaatt accatacata aactgtctca ttatacgtat gcaatttttt agttgtttt 2160
tgttttgtat aaat taattt gtttaattaa tatttcttaa gtataaacct tatgaactac 2220
agtggagcta cctcatttga aatgtaattt cagttctaaa aagatgtaat aactatttta 2280
gaattaaaa ttcttctact ttttaataaa ttatgaatat taaaggtgaa acctgtataa 2340
attactttga ttccatttta agtggagaca tatttcagtg atttttagta aactttaaaa 2400
atgtataatg acttttaaaa ttttgaagt tgaaggagcg ctaataaaaa tttattattt 2460

```

<210> 404

<211> 558

<212> DNA

<213> Homo sapiens

<300>

<308> Contig51464_RC

<400> 404

```

agagtttcaac aaataacatt tattgaccac attcataaca gaaactcacc cacattcaca 60
tgaactttgc tacaaccaag aatcaactga tgagaaatca tttgggttga atcctgggtg 120
tggattctgt tctgagaaga cagcgactgt cacagacgcc tctgtattgg tggactctca 180
aagattctga togtgacaca tttttgtggc tgttcttgac agtagctgaa gtttccaaact 240
tcaaacccca aacacactggc gaatggocac gatgaccctc gggatggcct gaggaggagt 300
caggaggccg aggggcagca gacgcgtgac agggagccgt agctggtcct ggttaggggc 360
ctggagctgc tggaaacttg acaacaggca ctcatcacg cagatccgcc agcagttgtg 420
ccagcagcat gcatccagca ggggtgctgc ttttggggtc tcaggagac agggaccag 480
taaggcagtg ctgcttccac cctggacaca ggcaggcgcc ccccgaaact ggtctggcat 540
tcacagcaga tggccagg

```

<210> 405

<211> 1929

<212> DNA
<213> Homo sapiens

<300>
<308> AL080079

<400> 405
gggggagaat ttccattagg caagctgtat gcagtggaat ttctctttaa ggagacacac 60
aattaaagact ctctgggtct gtcccttgctt tgggagacttt aagacatttc cttaaagacaa 120
aataaaaagcc tcgtatttcc ccattgagag ttttgttcca aggaatatga agtgagacat 180
atgggtgagt tataataatc aaaataattt atgaagagct ggggtctgcaa tagctagctc 240
aaaaactact tgtgtgtcag tctctctggt atagtatata agagcctgag gaagctctgc 300
aagatagatg gtglattatt tatggatcag gctgctgoat acaaacccttg catactatta 360
tcagcgttac ctaactctca gactattctg agtaatgctt gcttgctaat gaatgtatag 420
gagaccacat tgtaatgttt cttagatgat ggagtcacat cagtttctta gaaatcggtc 480
tcagtgcatg ctgtgctttt tcacatttgc tctgggttat ctgggaagta tcaggtctgc 540
ggaggaacaa gcattaaagt ataagaaaag gagacattct ggcaagcca atctgcttaa 600
agggcaagct cagaaacctg aacctagagg cctttctctc tgcaagaaa acaggtagtt 660
tgcatgctga gataggagg agcttttagg ctacacagca acccaaggga cctctacct 720
tttgcgtgag ttcaatcagg aagctatttg cctggctcca gcagatgatg agataatgag 780
gtagtggttt ttttattact gtccattttt gcaacatcct gcaacacct cctggagac 840
aagagcatta ccgagcttgg ctttcacggg ggaggggtgt attcagtaaa aaagaatagt 900
aaataaagg tcaactgagat tctaagtaag atagtaaatc taaggctcat gagccaatc 960
cttttcaata ggcataattt aacaggctgc ttaattttggc ggaggttaca tatggatgaa 1020
aatgaatcct agtcactgaa tattcatata catttccccc aaaaaccttg accatcatag 1080
tagatttttaa ttagttagct ttctaactag tcagattttc gcccaagtg ctttagtaat 1140
agtcattaaag atatagcttc aaaaatggct ttgcttttga tttctactca tttctgtatg 1200
ggtccagaaa aatatttttt tcatatttga caatgtcaca tccactttag aaattttcaa 1260
taaccagatg agaaaaaaat taagaaattg ctcaaggaaa acattttgaa atggatttga 1320
aagatttgagc caaattctgt tgtcagttct aagcatgcag ttctcacctc ccttagtcc 1380
cccatcagaa cagaggtcag gaatttagct ggggagccta aatttagtct agcttacctt 1440
tgagaatctg atcaattcag actctctttt ctattgtttt tctttctttt ttccctcttt 1500
ttaaactaca ttgtgtttag gtctgtatct aggatcctga gagattttcc atctctgtca 1560
ccattcacctt gcatgtttaa gattttcttt gtctgtgttt ggcatagatt cttttgtaca 1620
tatttatttta ttgtgtttaa tatattgtcaa ttggtttcct ttcttagctt gatattgctt 1680
agctttgttg ttttaattaa ctttctatta gagggactgt atataatttt tctaataact 1740
ttgtgaaact attttggta gcaatatctt tgaatatgat gaataaaagt gactgtgagt 1800
gcaaatagaa tttagcagtaa gaagctactc tagctaattt gccatttcac ttaaatggaa 1860
aatgtttttc aaataaatac ttatgtgttt catgttccaa gttattcaaa aaaaaaaaaa 1920
aaaaaaaaac 1929

<210> 406
<211> 3915
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006931

<400> 406
gtgggtggg gtggggctgg gggcttgcg ccttttcagg ctccaccctt tgcggagatt 60
ataaatagtc atgatcccgag cgagaccagc agatgcctgt aatggtgaga ctttggatcc 120
ttcctgagga cgtggagaaa actttctgct gagaaggaca ttttgaaagt tttgttgctc 180
gaaaaagctg tttctggaat caccctcaga tcttctttaga agacttgaat tagattacag 240
cgatggggac acagaaaggct accccaagctc tgatatttgc catcacagtt gctacacagt 300
gctctttcca atttggctac aacactgggg tcatcaatgc tctctgaaag atcataaagg 360
aatttatcaa taaaactttg accgacaagg gaaatgcccc accctctgag gctctgtcca 420
ggtcctctcg gtcccttgct gtggccatct ttctcgtgg gggatgatgc ggtcctcttt 480
cgctcgagct cttcgtcaac cgcttggcca gggcgaattt aatgctgatt gtcacacctg 540
tcgctgtcac tgggtgtcgc tttatgggac tgtgtaagt agctaagtc gttgaaatgc 600
tgatcctggg tcgcttggtt attggcctct tctcggagct ctgcacaggt tttgtgcccc 660

tgtacattgg	agagatctcg	octactggcc	tgcgggggtgc	ctttggcact	ctcaaccagc	720
tgggcattcg	tgttgaatt	ctggtggccc	agatctttgg	tctggaattc	atccttgggt	780
ctgaagagct	atggccgctg	ctactgggtt	ttaccatcct	tctgtctatc	ctcaaaagct	840
cagcccttcc	attttgcct	gaaagtccca	gatttttgc	cattaacaga	aaagaagagg	900
agaattctaa	gcagatcctc	cagcgggtgt	ggggccacc	ggatgtatcc	caagacatcc	960
aggagatgaa	agatgagagt	gcaaggatgt	cacaagaaaa	gcaagtcaac	gtgctagagc	1020
tctttagagt	tccagcgtac	gcacagccca	tcatcatttc	cattgtgctc	cagctctctc	1080
agcagctctc	tggtgcta	gctgtgttct	attactcaac	aggaatcttc	aaaggtcgag	1140
gtgttccaga	gcccatctat	gccaccatcg	gcgcgggtgt	ggttaatact	atcttcaact	1200
tagtttctct	attttctggt	gaaagggcag	gaagaaggac	tctgcatatg	ataggccttg	1260
agggatggc	tttttgttcc	acgctcatga	ctgtttcttt	gttattaaag	gataactata	1320
atgggatgag	ctttgtctgt	attgggggcta	tcttgggtct	tgtagccttc	tttgaattg	1380
gaccagggccc	catctccctg	tttattgtgg	cogaactctt	cagccagggc	ccccggccag	1440
ctcgagatgg	agtggcgggc	tgctccaaact	ggacctccaa	cttctcaacc	ggatgtctct	1500
tcctctccgc	tgctcaactat	ttaggagcct	acgtttttat	tatctcaacc	ggtcttccca	1560
tacacttctt	ggcttttaacc	ttcttcaaa	tccttgagac	ccgtggcgag	acttttgagg	1620
atatccacag	ggcctttgaa	ggcgaggcac	acggtgcaga	tagatctgga	aaggagcgcg	1680
catggagatg	gaacagcatc	gagcctgcta	aggagaccac	caccaatgtc	taagctgctg	1740
ctctctccac	ctcctcccg	gcattgggaa	gccactcttc	cctcaacaa	ggagagacac	1800
catcaggatg	aaaccaggac	gcttctgaat	gctgctactt	aattccttct	tcactccacc	1860
cacttccaga	gcaccccaag	gctgcgggtt	gttggatctt	caatggcttt	tttaatttta	1920
tttctggagc	atctcttct	gcttaggaga	gaccgagtg	acctaccttc	atttcaggag	1980
gatttgcg	cttggcacat	gaacaacttg	ccagcttttc	ctccttggg	ttctgatatt	2040
gcgcgactag	gggatatagg	agagaaaaag	taagggtgag	ttcccccaac	ctcagactta	2100
ccaggaaagc	gatacatatg	agtggtgaa	ccggagggtg	tttatgtaag	cagactcttc	2160
tcaactccat	acagctctac	gtggcaaat	aaactgagtt	ttatttttt	tatcctctg	2220
tttaattaca	taatttttt	tttttactt	taagtttcag	gcataatgtg	cogaatgtgc	2280
aggtttgta	cataggtata	tatatgccat	gatggaaata	tttattttt	taagctaat	2340
tttgcacaa	aataaaaaa	gaaggaat	gagattagag	ggaggtgtt	aaagagagg	2400
tatagatgag	aaagtttgat	gctggagagg	ttaaagtgca	ataaagaatt	agggaagaat	2460
gtgtgtcaat	attggagggt	aaatgatgtg	gtgcctgagg	tctgtacgtt	acctcttaac	2520
aatttctgtc	cttcagatgg	aaactcttta	actctcgt	aaagtcatat	acctatataa	2580
taagactact	gatttccctg	gagcttttt	ctttaagata	atagtttaca	tgtatagta	2640
cttgaatctc	aggattatta	actaatatgg	gaattgtagt	taatgatgtg	gtagggggtc	2700
taattttgga	tgaggtccag	ggaagagaaa	gtgattttca	gaaagcctgt	tcctccactc	2760
ggatgaaata	actcctctct	gtagtagtct	catctacttt	gaagtaatcc	cgccacatct	2820
ctcgtgggag	agccatccaa	ataaagaacc	taaaataatt	ggttcttgg	agagattcat	2880
tatttttcca	ctttgttctt	taggagattt	taggtgttga	tttctgtgt	tatttttaac	2940
catactttta	aaggaattcc	ccaagaatg	tttatagcaa	acttggaa	tgttaacctca	3000
gctctgggag	aggatttttt	tctgagcagt	tattctctaa	agtggtgtgt	gtcttaggc	3060
tcacgggca	cttgcgtatg	tctgttacc	gtcactgtg	gtcctatgoc	ggaatgcctc	3120
aggggacttg	aatccttcca	ataaaccag	tttagacagt	attagtaaat	gaatgctgta	3180
gcccacact	gagagatga	atgtatgtgc	actgtcactt	tgctctgggt	ggaagtacgt	3240
tattgttgac	ttattttctc	tggtttgtt	ctcacagccc	ctcagacccc	ttgtgtctag	3300
tctcccttcc	ctctcttgg	gcttacacat	tcctctgtct	ctcagacccc	cttgtctcgt	3360
gcagctatgt	tggtttctag	ttctcaactg	ctcagacccc	ctcagacccc	ttgaataaaa	3420
atgcacgttg	ctgagccg	atgcggtg	ctcagacccc	ctcagacccc	cttggagg	3480
ctcagccggg	cggtcagggg	ttcgagacca	gtctggccaa	catcgtgaaa	ccctgtctct	3540
actaaaaatg	caaaaattag	ccggcggtgg	tgccggcgcc	ctgtaatccc	agctacttgg	3600
gaagctgagg	cggggagaatc	atgtgaaccc	gggagcgagg	ggttgcaagt	agcgagatc	3660
gcacatgtgc	actctagcct	gggcccacag	gogagaactc	gtctcaaaaa	aaaaaaaagt	3720
cacatagctc	tcgagtgtgc	tttagcttga	aaaggtgacc	ttgcaacttc	atgtcaactt	3780
tctgctcctc	caaacagtag	gttggcagta	agggcgggtc	ccattttcca	ctgagagagt	3840
tgtgaatatt	tccatattga	ttttctattg	ttactctggt	tcttgttttt	aaaaataaaa	3900
ttctgaatgt	acacg					3915

<210> 407

<211> 2125

<212> DNA

<213> Homo sapiens

<300>

<308> AF257157

<400> 407

tgccgatggc	gtacttggct	tggagactgg	cgccgctgtc	gtgtccgagg	tcactagttt	60
cccggtagtt	gcactgcaca	tgaatagaac	agcaatgaga	gccagtcaag	aggactttga	120
aaattcaatt	aatcaagtga	aactcttgaa	aaaggatcca	ggaaacgaag	tgaagctaaa	180
actctacggc	ctataaagc	aggccactga	aggaccttgt	aacatgccca	aaccagtggt	240
atttgacttg	atcaacaagg	ccaatggga	cgcatggaa	gcccttgcca	gcctgccaaa	300
ggaagctgcc	aggcgaact	atgtggattt	ggtgtccagt	ttgactctt	catgtgaatc	360
ctctagtcag	gtggagcctg	gaacagacag	gaaatcaact	gggtttgaaa	ctctggtggt	420
gaactccgaa	gatggcatca	aaaagatcat	gttcaacccg	cccaaaaaga	aaaatgccat	480
acacactgag	atgtatcatg	aaattatgct	tgcaattaaa	gctgccagca	aggatgactc	540
aatcatcact	gtttttacag	gaaatgggtga	ctattacagt	agtgggaatg	atctgactaa	600
cttcactgat	attccccctg	gtggagttaga	ggagaaagct	aaaaataatg	ccgttttact	660
gaggttaagag	ttactgcact	ttgtctgatc	ccagagggcc	cggtcttagt	catcaaatgt	720
gtatttagtt	gagcaaaagt	gcaatatctg	cttctctgtg	tcagtgcagt	taattgtgac	780
tttgacaacg	catatttctc	taaaactggct	ccatccatag	gtgggcatcc	attctttctc	840
tttgacctgt	gttacctgtt	catcctgctg	gtctcatgag	atttgactgg	aaaacataaa	900
atgtgactat	agttgttaac	aaaattatct	tgcatgtagg	cagaggaatt	tagcaataat	960
tcactccagcc	agaggaaatg	gagatttcca	tgaacgcgtg	atcgaggaga	ctggaggagg	1020

tcttgaatac	acaggaggag	gtcatgaagg	cttttctctt	tattctgccc	aaatctgtaa	1080
ttcccaactg	ctgtgcatatt	gcagcatgtg	ctcagtcacc	caacctaatc	tctaagactg	1140
cggtctgttt	cgagccctga	actcaactgtc	cagaatacaa	gattcaaac	agggttgacc	1200
cataggtctc	actggttggt	cttccatcca	ctgccattca	gacggatgcc	ctgtagaat	1260
ctggtgcgga	agaccaggga	agtggcaggga	caaggaggca	ctgctgggtg	tggtctgtga	1320
aaaccccgag	ctgacagaag	gctttatctc	agagtccagc	tgcttgcgcc	tttcttcagt	1380
catgctgctt	tgccacattt	aatgatattg	tgtttccatt	ctgttaggga	attgtggggc	1440
tgttttatag	attttcttaa	gcctctgatt	gcagtggtca	attggtccag	tggtgggcatc	1500
tcctgtcacc	tccttgggct	attcgtagcc	gtgtatgcac	ctgacagggc	aaactttcat	1560
acaccattta	gtcacctagg	ccaaagtccg	gaaggatgct	ccctttacac	ttttccgaag	1620
ataatgagcc	cagccaaggc	aacagagatg	cttatttttg	gaaagaagtt	aacagccggga	1680
gagggatgtg	ctcaaggact	tgttactgaa	gttttccctg	atagcacatt	tcagaaagaa	1740
gtctggaccca	ggctgaaggc	atttgcaaat	cttcccccaa	atgtcttgag	aatttcaaaa	1800
gaggtaatca	gagaaaagga	gagagaaaaa	ctacacgctg	ttaatgctga	agaattgcaat	1860
gtccttcagc	gaagatggct	atcagatgaa	tgacacaaat	ctgtggtgaa	cttcttatcc	1920
agaaaaatca	aactgtgatg	accactacag	cagagttaag	catgtccaag	agagcatgtg	1980
ctgttacctc	tgattttccag	tactggaact	aaataagctt	catgtgtcct	tttgtagtgc	2040
tagaatatca	attacaatga	tgatatttca	ctacagctct	gatgaataaa	aagttttgtg	2100
aacccaaaaa	aaaaaaaaaa	aaaaa				2125

<210> 408

<211> 1446

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002916

<400> 408

ctgccattta	ggacaagctg	gatgatgatg	gtttgatagc	tccagggggt	cggtgatagg	60
agatgatgaa	tctgcttcat	ccagaatcac	aatcttaaaa	ggcgggaact	gaggcgactg	120
tggggacatc	agtgatcgta	agtctcctgg	gcccgattat	ctcagattag	gtgacggagc	180
taagactctg	agaccatctc	gtcctttttg	tatagcggaa	acctgaggaa	cgagccggcg	240
cggtgacctc	gcacgagaag	ccaggctaac	tggtggaagt	accatgcga	catcttttaa	300
aggtacattc	atcatacta	aaccoccgct	gaccaaggat	cgaggagtga	ctgcccagtc	360
gggaagttag	ggagagaaca	agaaagccaa	accogttccc	tggttggaag	aatatcgccc	420
aaaaatgtg	gatgaagttg	ctttccaggga	agaagtgggt	cgagtgtgta	aaaaactctt	480
agaaggagca	gatcttccca	atctcttggt	ttaaggacca	ccctggaactg	gaaaaacatc	540
cactattttg	gcagcagcta	gagaactctt	tggtgctgaa	cttttccgat	taagagttct	600

tgagtgtaaat	gcactctgatg	aacgtggaat	acaagtagtt	cgagagaaag	tgaataattt	660
tgtctcaatta	actgtgtcag	gaagtcgctc	agatgggaag	ccgtgtccgc	cttttaagat	720
tgtgattctg	gatgaagcag	attctatgac	ctcagctgct	caggcagcct	taagacgtac	780
catgggagaag	gatgcgaaaa	ccaccgcgatt	ctgtcttctc	tgttaactatg	tcagtcgaat	840
aattgaaccc	ctgacctcta	gatgttcaaa	attccgcttc	aagcctctgt	cagataaaat	900
tcaacagcag	cgattactag	acattgccaa	gaaggaaaaat	gtcaaaat	gtgatgaggg	960
aatagtcttat	cttgtttaaag	tgtcagaag	agacttaaga	aaagccatta	catttcttca	1020
aagcgtctact	cgatttaacag	gtggaaagga	gatcacagag	aaagtgtatta	cagacattgc	1080
tggggttaata	ccagctcgaga	aaattgatgg	agtattttgct	gcctgtcaga	gtgctctttt	1140
tgacaaacta	gaagctgtgg	tcaaggattt	aatagatgag	gtctatcgag	ctcagcagt	1200
cgtaactcaa	ctccatgatg	tgtttgtaga	aaataaacta	tctgtataaac	agaagtctat	1260
tatcacagaa	aaacttgcgc	aagttgacaa	atgcctagca	gatggtgtgt	atgaacattt	1320
gcacactcatc	agcctttgtg	caactgtgat	gcagcagtta	tctcagaatt	gttaacgtga	1380
tatatctgga	tggtgggttt	tgtaaataat	gaagttgttaa	taaaaataaa	atgacacaaa	1440
gcacgcg						1446

<210> 409
 <211> 788
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig55725_RC

<400> 409						
caagttagaa	caaacatttt	attattaaaa	taaacttttg	tataaaagca	ttacagatca	60
aaagctgtat	ttacactttat	cgattcaagg	tccaattatg	catcaaacat	tgaattgcac	120
agcaatgggt	tacatatgca	agtaaatgtg	acatacaaac	acttagattc	cacctctacc	180
aaataccttg	attaatgcga	agaggagggg	gaatactgac	acaggaaaac	tgcccagaaa	240
ctagacttgc	agagatgtca	ggttaacaaa	ctgtctaaag	ttacatctcc	aaaaaggcac	300
ttatcattgt	tataaaagtg	cttaaaactc	aaacttgaa	cttgtgcctg	gtttataaat	360
ttacaagaaa	ctgcacagaa	ccacagacta	gtttttaata	tcaagtttcc	atacaaaatt	420
gtccaagaa	ttatttgcga	taccttacat	gtgaactgaa	ataaactttg	caaacacaaa	480
ttatactaaa	gtttttatact	gaaaattcaa	atcagcttac	tcataatctc	cagatagctc	540
ttctcatcatt	atttttaatta	agtattacat	tgcccaccaa	cctttgaatt	atgatcgtgt	600
aagctggacg	caccacttac	ccttcgggca	gaagaaaatta	ggagaaaatta	agtcctcaact	660
gtgacaacaa	cttgtaacat	aagcgcgaata	catgggattg	agagacatca	atctgcaata	720
attacctaat	ttttccacat	ggttttaaga	cttttgtgcc	ttgattctat	taaaaatttc	780
aactagta						788

<210> 410
 <211> 433
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig24252_RC

<400> 410						
cagctaacac	aaaagctttt	acttggaac	tggcaaatat	tggactagaa	tactgacatg	60
ctcagctctc	ggcggagctg	cgatgcaga	agctatttgc	caagccctc	ctgattgcct	120
tgtctctctt	tggcatgtat	cagcagagcc	ccaaggccca	attgccacca	gggtgggagc	180
gttctccatc	aaggtatggg	gaccctact	ctcttgtttt	gttaaaaaag	gcaggtaggg	240
gaagaagccc	aggcagttga	cccagctctt	gaacagctgt	actccccagc	accagtgatc	300
cactggcaga	tgcagggttt	gcacctatc	cagccctgtg	ttccttggtt	gggctcaagt	360
tgtcagtgtc	caggtacctg	ggctgggcca	gagctggggc	caggggagcc	tgtgtttgag	420
gggtaagggc	tgc					433

<210> 411
 <211> 969
 <212> DNA

<213> Homo sapiens

<300>

<308> AF201951

<400> 411

```

acggaacacgc  cagtgaggagc  ttccagctga  gcgctcccca  gaggtgagct  gatccccagc  60
cacgacacac  aggcacagc  tgcgagaa  gcacatcag  catcatgcta  ttacaatccc  120
aaaccatggg  ggtttctcac  agctttacac  caaaggcagc  cactatccct  caaagagaga  180
aacctggaca  catgtaccac  aacgaagatt  acctgcagaa  cgggctgcc  acagaacaca  240
ccgtctcttg  gactgtccag  atcctgtgtt  gcctgtgat  ttcaagtctg  gggggccatct  300
tggtttttgc  tccctacccc  tcccaactca  atccagcaat  ttcccaacct  ttgatgtctg  360
ggtacccatt  tttaggagct  ctgtgttttt  gcattactgg  atccctctca  attatctctg  420
gaaacaacat  aactaaagcc  tttagctga  gcagcttgac  ctcaaatgca  gtgagttctg  480
ttactgcagg  agcagggctc  ttccctctg  ctgacagcat  ggtagccctg  aggactgcct  540
ctcaacatgt  tgggtcagaa  atggattatc  tatccctatt  gccttatctg  gagtactatt  600
tccaatata  tgaaatcaaa  gattgtctcc  tgaccagtg  cagtttaaca  ggtgtcctag  660
tggtgatgtc  catcttcact  gtgtcggagc  tcttattagc  tgcatacagt  tctgtctttt  720
ggtggaacac  gctctactcc  aacaacctg  ggaggtcatt  ttctctgacc  cagtcaaacg  780
atcatatcca  acaggtcaaa  aagagttctt  cacggtcttg  gatataagta  actcttggtc  840
tcagaggga  gaaaagcaac  tcaacactca  tggtaagtgt  tgattagact  ttcttgaat  900
ctctgccatt  ttagatactg  tgaacaaac  taacaaaaaa  aaagcttttg  tttgttaaaa  960
aaaaaaaaa  969

```

<210> 412

<211> 2917

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005915

<400> 412

```

ggaaaaaagc  gacttgtggc  ggtcgagcgt  ggcgcaggcg  aatcctcgcc  actaagcaaa  60
tatggacctc  gcggcgccag  cggagccggg  gcgcggcagc  cagcacctgg  aggtccgcga  120
cgaggtggcc  gagaagtgcc  agaaactggt  cctggacttc  ttggaggagt  ttcagagcag  180
cgatggagaa  attaaatact  tgcaattagc  agaggaaactg  attcgtcctg  agagaacac  240
attggtttgt  agttttgtgg  acctggaaca  atttaaccag  caactttcca  ccaccattca  300
agaggagttc  tatagagtgt  accttacct  gtgtcgggcc  ttgaaacat  tcgtcaaaga  360
ccgtaaagag  atccctcttg  ccaaggattt  ttatgttgca  ttccaagacc  tgcctaccag  420
acacaagatt  cgagagctca  cctcatccag  aattggtttg  ctcaactcga  tcagtgggca  480
ggttggtcgg  actcaaccgc  ttcaaccaga  gcttgtgagc  ggaacttttc  tgtgcttggg  540
ctgtcagaca  gtgatcagg  atgtagaaca  gcagttcaaa  tacacacagc  caaacatctg  600
ccgaatacca  gtttgtgcca  acaggaggag  attcttactg  gatacaaat  aatcaagatt  660
tgttgatttt  caaaaaggtc  gtattcaaga  gacccaaagt  gagcttccct  gaggagat  720
ccccgcagtc  tttagagtaa  ttttaaggcc  tgaagctgtg  gaatcagctc  aagctggtga  780
caagtgtgac  tttaaggga  cactgattgt  tgtgctgac  gtctccaaag  tttagcaccc  840
aggagccagt  gcagaaacta  attccctgt  cagtgtgtt  gatggatatg  agacagaagg  900
ctctcgagaa  ctccgggcc  ttggtgttag  ggaacctttc  tataggctgg  ctttctctgc  960
ctgctgtgtt  gcgccaaaca  accaaaggtt  tggggggaaa  gagctcagag  atgagggaac  1020
gacagctgag  agcattaa  accaaatgac  tgggaaagaa  tgggagaagc  tgtttgagat  1080
gagtcagaat  aaaaatctat  accacaatct  ttgtaccagc  ctgttcccta  ctatacatgg  1140
caatgatgaa  gtataacggg  gtgtcctgct  gatgctctt  ggtggcgtct  caaagacaac  1200
aggagaaggg  acctctcttc  gaggggagat  aaatgtttgc  attgttggtg  acccaagtag  1260
agctaaagac  caatttctca  agcacgtgga  ggagttcagc  cccagagctg  tctcaacagc  1320
tggtaaagcg  tccagtgctg  ctggcttaac  agcagctgtt  gtgagagatg  aagaatctca  1380
tgagtttgtc  attgagctg  gagctttgat  gttgctgat  aatggtgtgt  tgtgatgta  1440
tgaatttgat  aagatggagc  tgcgggatca  agttgctatt  catgaagcta  tggaaagcta  1500
gaccattatc  atcactaaag  caggagtgaa  ggctactctg  aacgcccggc  cgtctatttt  1560
ggcagcagca  aacccaatca  gtggacacta  tgacagatca  aaactattga  aacgaatat  1620
aaatttgtca  gctcccatca  tgtcccgatt  cgatctcttc  ttatcctctg  tggatgaatg  1680

```


ttgtctgaga	agccactgat	ctctgaggag	acggacotta	ttgagccaac	tctgtcggat	1860
gagctaactc	gccacaactgg	ttctttggcc	tctgtgtatc	ataagcctcc	caatgctttt	1920
gtgagaagaa	gtcatggaat	tcactcgtaaa	cacttgccaa	ttcatcatgg	gagcactgat	1980
gcaggtgaca	gcctctgttg	cactaccact	gcaacgaacc	tggaaacagcc	tcaggttatc	2040
ccctctcaag	gtgatcttct	aggggatctt	ttaaaccctg	acctcggctcc	cccagtgcaat	2100
gtgccacagg	gtgtccctcat	gcagatggga	cactgtggat	ttcctaggagg	aggtactagat	2160
agtcctgttg	gacaatccct	catcccatca	tcggtgctgt	caacctttgc	tccttcaact	2220
acacctgctg	ttgttcagcag	tggactgaat	gacctgtttg	gacctctcac	agggatagcc	2280
atggcacctg	gtggtatgtg	ggctcctaag	gctgtctggc	tacctgcagt	aaaggctaaa	2340
ggctgtgaga	tttccggaaac	atttaactcac	cgccaaaggcc	acatctatat	ggaatagacc	2400
ttcaccataa	aagctctgca	gcacatgaca	gattttgcaa	tcacgtttaa	caaaaatagc	2460
tttgggtgtc	tcgccagcac	tcctctggcc	atccatacac	cactgatgcc	aaaccagagc	2520
attgatgtct	cctctgctct	caataccttg	ggcccagctc	tgaagatgga	acctctgaat	2580
aaocctcagg	tggtctgtgaa	aaacaatatc	gatgtctctt	acttcagctg	ctctacccca	2640
ctcaatgtgc	ttttttgtaga	agatggcaca	atggagcgcc	aggtcttctc	tgcacatagc	2700
aaggaatatt	ccaatgaaaa	tgaacttcag	tttcagatta	aggaatgtca	tttaaatgct	2760
gacactgttt	ccaagcaagt	gcaaaacaa	aatgtttata	ctattgccaa	gaggaatgtg	2820
gaaggcagg	acatgctgta	ccaatccctg	aaactcacta	atggcatttg	gattttggcc	2880
gaactacgtg	tcacagcagg	aaaccccaat	tacacgctgt	cactgaagtg	tagagctctc	2940
gaagctctct	aatacatcta	tcaggtctac	gacacgatct	tgaaaaaacta	acaagactgg	3000
tcagctaccc	ttcaaccatg	ctgtgatcgg	tgcgaagtc	gaactcttaa	ctggaaagaa	3060
ttgtattgtc	gcgtagaatc	tgaacacact	gaggccacct	agcaaggttg	taactagtct	3120
aaocctgtct	acactatagg	cacaacctgt	ttagcttagtt	ttagcttcttg	tgaacatttg	3180
taaccactgc	ttcagtcacc	tcaccaactc	tgcacactgc	tgtcgtctatc	tgtccttact	3240
tgtgggtctc	tcacatctgt	gccaatggct	ggctttgtct	acacctctct	tttagttagt	3300
tttgttattt	tgtaatgtag	agctcatctc	aaaagcagaa	aaagacaaca	aatatgaag	3360
caaggaaaag	tgtcaactgaa	acactgcact	ttactgtttt	atacttttgt	acatatgaag	3420
aatcaaggga	ttagtcaaac	cagttagaag	cattgaaatg	actgtcatta	accacaagc	3480
cotgaggcca	gagatgcagt	tacotacact	agctttttgt	gggttctctt	acctgtagta	3540
gccttatccc	ttgtctattg	gattttcagt	ttgttttttt	ctttttttcc	ccctccaaact	3600
ccctttctct	ggccaagcct	tcactgtctc	ccctttccat	attataactc	catttgattg	3660
ctctgcagtt	gggaacggtg	atcttcttga	agtgctgttc	agtatgataa	aatcatagag	3720
ctctgcagca	ccaaagctga	cagaagttat	acottactcc	tttctttccc	cctgaacaaa	3780
ccctgtaact	ccactaatct	aggaatttga	gtagagatgg	ggaaacaagaa	ccagatgtct	3840
gtccctccac	ccctctctct	gtattttctc	ggtccagttg	aaatctaaaa	ttctactttt	3900
agagttgaaa	cagagtaata	acttatctaa	ccctcttttc	ctacaaaagg	gaaagataaa	3960
agggcacaag	gttaacgcc	agggccgtca	gctgtgtagt	ggcaaaagccg	agacccagtc	4020
tcctaagtcc	ccgtctagtgt	ggttttccacc	acaggactgt	ctcttgtctgt	tttccctaaa	4080
tgccctctcc	tgcccttttct	gtgcctagtt	tttggtctct	cacatatccc	atattgattt	4140
tgaactctct	tatatattgca	tcaggtggca	ctcgaatate	ttttgaatta	ctcgaaggtt	4200
aagccagatg	ccagaatgaa	ggtgtagcca	gtgtttccca	tatgccctct	gagccccaact	4260
tattgaggcc	agcagaatag	gtgcagagat	gtcagtagct	gaaagttagt	gcmaaagtct	4320
tttatccctt	cagctctctg	atctgctctt	ttctcatgat	acttagtctg	caggggcaat	4380
taagatcate	ccagaagttc	agggcagttc	tctctactct	gaaaagactg	ggggatagta	4440
aatcttcccc	ctaccccact	taatgcgttg	gatattgatt	ttcaaaagat	ctgtcatgac	4500
caaaatacca	gcctgttttag	cagtgttaca	ctgttttgat	tgcggggcact	tggtgcattg	4560
cctggcacc	aatattcagg	gtccatgact	aaagctggtc	ttctcagatg	ccctgtctaa	4620
atcaggggca	cttcaggctc	cacaggcgct	atgttggact	gagacctaac	tcactggact	4680
cagaggagga	atcgtggaaa	acaagagcca	aaactaccca	cccccctatt	tcagtctgta	4740
aatcaacctg	tttcatacca	gttgcacagg	ttgtggggag	cggtcccca	aagcaacttt	4800
ttaaaccttg	agaaactcca	agagaaaaat	atttggggaa	ggaggaggga	aatatgtccc	4860
ttgacacaca	ccctcggaag	acatggcagt	aggaacacgc	ataggatttg	atgtggggag	4920
tggataggct	ggtgatgtgt	ggagcggaaa	agcaggtttg	taaaagtccc	ttcttgggac	4980
ttattcctgt	agtcagtgga	tacaagttag	gcagaaggtt	cacactgcaa	atagtgttct	5040
catctcaaa	caaacactca	ttccagaagg	aaaagtgtgt	gagcccaagc	agacacaaca	5100
atttctcatc	agaatatgtc	cctcaacccc	cgaacaaagg	ctctctccag	ctctccccc	5160
agtgtggat	aacagactcct	attctcagct	gactgactgt	agccaaccca	tgaactcttc	5220
actccttggt	gaagccacct	cccatcacac	ccctgagcgc	agtttagggag	gaattctact	5280
ttccataaaa	ggacctctcc	tgagaggcaa	aaactgttgc	ctccaccagc	gcttccctct	5340
tggctcatct	caagcttggt	caaatgggg	aagtgggatg	gaggttgccc	tgcctcccc	5400
ctcctctgcc	tgagtgtgtc	tttgaatagt	cagctggcat	catacaagaa	gcaggagag	5460

```

caaacaccca gaactctttt gctggtcaga gattccctga gtgtctgtcc tcacccaagc 5520
ctgctctgtg tctgtgttgt gaagcttgag actctggaaa gaaatgggga gggggggcag 5580
gggaaatgtt gccttaagaa tgctttccat tctctgtgtc ttattgggtc ctgtttttcg 5640
ggagggtggg ggttggggga agcttgacct tgtgtcttcg tcaataaact cacatttaca 5700
c

```

<210> 414
 <211> 1165
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56457_RC

```

<400> 414
gcaagtcttt ttttttttt taataaaacc aacataacag aatgtagaaa tctaaccgaa 60
ctattctccc tcaaaagtgt tacatcctta gctccacaaa aactactatt gtggagagct 120
tatttagaac tctctgttga ggaagggcct taaaagggtg tttatgagcc acattagaaa 180
atacccttct aaactttggg ctggttataa acagaaaatc catactcaga ggaatgacac 240
gagaaaatla aaaaggtgct gcagctctca cggagggtcc ctaagaacct ctagcaagtt 300
ctgagcaaca gcatacccat tggagaaaac gttagcctcc caggacagcc actttccatg 360
ggacggcatt catttgata gaaaagttct gctgtctgtt taaaaagaaa aaagaaaaaa 420
aaaaaatcag ccacactcct ttatagtcac acttgcatat gattaaaaaa aaaaacaaca 480
acaaacacac tcacacgtgg ttcttgtatg aattctaat aagagacatt atattcaaga 540
ttatagcact catagtata attttcagaa gactataaca gcttctttgc cacttaaggg 600
gtacagaggg tgcagctgtt tacctgaatg aattcttaaa taagtgaact ctaccactct 660
aattttccta gttgtgagtc tgttacaggt tactgacgcg caagtccatg ttccagacat 720
tctttgccaa cacagattaa ttacaagaga gccaccccca aagctctagg gcgaattggg 780
aaaggggcgg ttaaatgagg aaacaggcca ttctcttgta aaagtttgta gttttcttcc 840
ccaatacact tcttttttat tgaagaacttc catttttttc ctgtgtccct cctctgcact 900
ctcacactct cgctctatct gctggaattt ggccacatgc acctcaatca ttgtatacg 960
ttgcagttgt aagatggaca ttgtgtgatg tgtgtcttcc tcatgcactc attttaaagc 1020
acccttcotga gaggttttgc tctctgagtt tttattggct gttattctga cagctgaaca 1080
cgagggttca ggttgagagg tcccctcaca tactggaaaa tgtatgaact ctttctcatc 1140
atcgacagtt tctctatttg aaaca

```

<210> 415
 <211> 867
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_020386

```

<400> 415
cgtttcagcg tggcgccgct ggtgctggcg ttggccctgg aggacggccc cgagtgatgg 60
ctggcgccgtg tctcccggtt gctctccggg tacagatgga gtctcccggt ggcgcgcggc 120
gcgaaggtcg gcagctcgga gcccaagaga gacccaggga cacacacagc gtctcccggt 180
tgcgagaaga agaccccggc ttgagagtga gatggcgttc aatgattgct tcaagttgaa 240
ctacccctggc aaccctgcc cagggggactt gatcgaagtg ttccgtcctg gctatcagca 300
ctgggcccctg tactttgggt atggttacgt tcatcaacata gcacctgata atggcaatcc 360
tgcgtctctt acaagcgcca agtctgtatt cagcagtaag gccctgggta aaatgcagct 420
cttgaagatt gtgttgaggaa atgacacata cagaataaac aataaatcag atgaacacta 480
ccccctctc cctgtggaag aaatcataaa gccgtcagag tttgtaatg gacaggaggt 540
ggcctataac ttactttgtca acaactgtga acattttgt acattgcttc gctatggaga 600
aggagtttca gagcaggcca accgagcgat aagtaccgtt gaagtttgtg cagctgctgt 660
tgggtctctc tcattctctg gcttgtttcc aaaaggacaa agagcaaaat actattaaca 720
atttaccaaa gatatattga tattgaagga atttggagg aggaaaaagaa acctgggggt 780
aatacttatt ttcagtgcat cattactgtt ccagattctc atgatggatg gcagactctt 840

```

taataaaattg cttactgata ttatctt

867

<210> 416
 <211> 3432
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> 2741
 <223> n = A,T,C or G

<300>
 <308> NM_014889

<400> 416
 ctgcctcatgc gcaatgtggc gctgcggcgg ggcggcaggcg ctgtgtgtgc tgaggcggct 60
 gagcgccgga catgcacacc acagagcgtg gcgatggaa acgtaacggg cttgtgagag 120
 ggctctgcag tataaactag gagacaagat ccatggattc accgtaaacc aggtgacatc 180
 tgttcccgag ctgttctcga ctgcagtgaa gctcaccat gatgacacag gagccaggta 240
 tttaaccctg gccagagaag acacgaataa tctgttcagg gtgcagttcc gtaccactcc 300
 catggacagt actggtgttc ctccacttct tgagcatacc gtcccttctg ggtctcaaaa 360
 atatccgtg agaaaacctt tcttcaaaat gttgaaccgg tccctctcca cgttcatgaa 420
 ccgcttcaca ctctagtgatt atactctgta tccattttcc acacaaaatc ccaaggactt 480
 tcagaatctc ctctcggtgt atttggatgc caccttttcc ccatgtttac gcgagctgga 540
 ttcttggcag gaaggtatggc ggctggaaca tgagaatccg agcgaccccc agacgacctt 600
 ggcttttaaa ggaagtgtct ttaatgagat gaagggagcg ttacagaca atgagaggat 660
 attctccagg caccttcaga acagacttct tcccgaccac acgtactcag ttggtctccg 720
 ggggtgacca ctgtgcactcc cggagcttac atgggagcag cttaaagcgt ttcattgccac 780
 tcactatcac ccaagcaatg ctagggtctt caagtacgg aattttccat tagaacagca 840
 tctgaaccaa attccagagg aagcactgag caaattccag aaaattgaac caagcacctg 900
 ggtgccagct cagacacctc gggacaagcc tagggaattc cagataacat gtgcccggga 960
 ttcatttgct acagatccct ctaaacaaac aacgtcagc gttagcttcc tcttaccgga 1020
 catcccgac acatttgaag ccttcacatt aagtcctctg tcttcaactc tgaattcttg 1080
 gcccaattct cctttttaca aagccttgat tgaattctgc cttggcacag aattttctcc 1140
 tgaatttggg tataatggct acacgaggga ggctacttt agtgcggccc tccaagggtc 1200
 tggggagaaa gacattgaga ccgtcagaag cctcatagac agaagcattg atgaagtgt 1260
 tgagacaagg attgaagatg atcgaattga ggctttactt cataaaattg aaatacagat 1320
 gaaacatcag tctaccagct ttgggctgat gctgacatca tacatagctt cttgctggaa 1380
 ccatgatggg gaccctgtgg agctcttgaa ttggggaaaat cagttagctta aatcagaca 1440
 tggcctgcag ctgaactccaa aatttttgca agaaaaagta aaacagttat ttaagaataa 1500
 ccagcataag ctagctttat cgtagaggcc agatgacaag tatcacgaga acgaggcaca 1560
 ggtggaagcc acgaagctca agcagaaggt cgaggtctgt tcccccgag acaggcagca 1620
 gatctcagag aaaaagctcag aattacggag tcaacaaagc aaactccaag atgcctcttg 1680
 tctgccagcg ttgaaagtgt ccatatttga acccaccata cctgtcacag agtctggacgt 1740
 ggtcctgaca gctggagata tccctgttca gtactgcgcc cagcccacca atgtgactgt 1800
 gatttccgg gccctctcca gctgaacac actccccgag gagctgaggc cctatgtgctc 1860
 cctcttctgc agcatctcca ccaagctggg ctgcggcctt cttgactacc gggagcagcg 1920
 tcagcagata gaattgaaga ccggagggtg gagtgtctct ccccaagtcg tcccgcagca 1980
 ctcacacatg gacacctacg agcaggttagg tgtgtcttct tctctctctc gccctgagct 2040
 aaactctgca gaactgatgc agctatggag tgaatatatt aacaaacctg gctttgaaga 2100
 agaggagcac ttaacaggtc ttgtgaagat gaccgccag gagctgcaca atggaattcc 2160
 ggaactctgg cactctgacg catccatcag accctcacg accctcacg ccgcagggga 2220
 cctgcaggag accttcacg ggaaggatca ggtgcggctg atgaagagga ttgcagaaat 2280
 cagagatc aaacccatcc tgaggaaagt cccgcgtatc aagaacact tgttaaatgg 2340
 tgataaatg aggtgttcag tgaatgcgac tccatcagca atgcctcaga cagaaaaagc 2400
 ggtcgaagac ttccttagaa gcatcggtgc gactaaaaa gaacggagcg ctgtgcgccc 2460
 acacacggtc gagaacctg tgcaccagc cttctgttga gatgccacg tccccattgg 2520
 ctccaggtc attaggaagc ttgtcatgga acccaacctc aagccctggc agatgaagac 2580
 tcaattctgt atgccttccc cgtggaatta tgcatccgaa ctgtccccca 2640
 cacggacca gatcatgca gttctaaaaa ccttgcagct ttgatgactg ccaaatctct 2700

gcatacagaa attcgagaaa aaggcgggtgc ttatggtgga ngcgcaaaa tcagccacaa 2760

tgggattttc accctttact cttacaggga cccaaataca atagagacgc tccagtcttt 2820
 tgggaaggct gtgcactggg ctaagtctgg aaaattcaca cagcagaagc tcgacgaagc 2880
 caaactttct gtcttctcaa cgttagatgc tctgtcgct cctcagaca aaggaatgga 2940
 ccaactcttg tacggcctct cggatgagat gaaggcagcc cacagagagc agctctttgc 3000
 tgtcagccac gacaagctcc tggccgtgag cgtatggtag ctccggcactg ggaagagcac 3060
 acagggcctg gccatctcgg gaccogagaa cccgaaaatt gccaaaggac catcctggat 3120
 catccgatga gcacgcgtgg cgtctgactg cacaggagcc cgagacaata cacctccaag 3180
 ctgaatatga aaagtacagaa atgtctactg tttttccaa aatattatgt cattgtgtgt 3240
 cgccaaagcc cttgactggc gagtcaaaaa ctgagatcta tcttaagagt gaccaggaag 3300
 aggttcattg aaataatcat gcatgaagcg ccaagatgc accatgtaga atttcaact 3360
 tgtactggca ggctcgtttt acctgattct agaataatta agaactcaaa aataaagggc 3420
 aactctgact ta 3432

<210> 417

<211> 817

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006681

<400> 417

agtctcgctg cggggccccc aggcgcagca gggcaccagg tggagcacca gctacgcgtg 60
 gcgcagcgca ggcctccctag caccagacct cccgcagcgg ccgagatgct gcgaacagag 120
 agctgccgcc ccagctcgccc cgccggacag gtggcccgcg cgtcccgctg cctgctgctg 180
 ctgctgctcg tcgctgggtg cgcggggccc tggcgagggt ctccaatatt acctcaagga 240
 ttacagcctg aacaaacagc acagtgttgg aatgagatag atgatacttg ttgctctttt 300
 ctgtccattg attctcagcc tcaggcatcc aacgcactgg agggagcttg ctttatgatt 360
 atgggaatgc taccaaagcc tcaggaaacaa gatgaaaaag ataatactaa aaggttctta 420
 ttccattatt cgaagacaca gaagtgtggc aagtcaaatg ttgtgtcgtc agttgtgcat 480
 cggttgtgct agctcgttcc tcacctgcat gagagaagaa tgaagagatt cagatggagc 540
 gaagaattcc aaagtccctt tgcaagtcaa agtcgaggat attttttatt ccaggccagg 600
 aatggaagaa ggtcagcaagg gtctcatttaa aatggatgcc agctaaattt ccaacagaga 660
 atgctatgga atacaaaatg tactgacatt ttgttttctt ctgaaaaaaa tctctgtcta 720
 atgtactctg ttgaaaatcc ctgtgtgtgc aatgttctca gttgtacaaa tgttgtaat 780
 gttcaatttg ttgaaatta aaaaatctaa aaataaa 817

<210> 418

<211> 5210

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007203

<400> 418

ccagaggctg agccccggcg agccccgcgt gcgcacagct ctgcccgccg cccctggagc 60
 ggaatccccg ccgtccctgg gcgttctccc ggggtctagc aaagccccc cgggtgcaca 120
 tcagtttctc tgggtgaacta cagcgtgtgt tttttcttcc ctctttcccc tgccgtgtgt 180
 ccctctctca ggtatgcaga ggcggaattg cacaaggaaa ggctgcaagc catagacaaa 240
 aaaaagaaag ggcagactga aatagaaggc aagcgacaac agcttgaaga gctgtgact 300
 ctgctgcagc attccaagtc caaagtctct cgggagaaat ggctgctgca gggcataccc 360
 gctggaactg ccgaagagga ggaagccagg aggcgcagct ctgaagaga ctgattcaga 420
 gtaagcaaac ttgaagataa cattcagagg ctggagcaag aaatacaaac cgtagaagat 480
 gaagagctcc agatatctgc caaagagcaa atcatctag agaaactgaa ggaacagaaa 540
 aaatctctca aggactttca gaagggtttc tccagtcagg atggagatgc agtaaatatc 600
 atttctcccc agcttcccca cctgccaatc ctctgttcac gaacagcaga acctcaact 660
 gggcagagac ggaccagcag agcgcgtgga gtcgggtggg agaattgtct gctaaaggaa 720
 ggtgagtcag cctcgaaagc cacagaaaca tccggcccag acatgactat caagaagcct 780

ccccagcttt	ctgaggatga	tatctggcta	aaaagcgagg	gagacaacta	tagtgccacc	840
ctctctggagg	ctgtgtgcoag	ctctcttttcc	ccagatocaca	aaaacatgga	aattgaggg	900
tcgtgttcgac	taatgtaaaag	tgttctctgga	atcacctcta	ccccacatcc	catggaccat	960
ccctccgcgtt	tctattccacc	cccgcataat	ggcctcotta	ctgatcacca	cgaatccctg	1020
gataatgatg	ttgocagaga	gatccgctat	ctagatgagg	tgctagaggc	caactctggt	1080
gattctcgtg	tggatggaaac	gtacaatgga	acatcctccc	cagagcctgc	tcagtggtgt	1140
ctggtggggcg	gcctaaagccc	ccctgtccac	gaggcgaccc	agccagaacc	cactgaaga	1200
acagctacgac	ggcagcgacc	tctccacatc	gagctcagta	atagcagccc	tgacccgtgc	1260
gcagaggccag	aaagaacaaa	tggccattcc	cccagccagc	ctagagatgc	gctgvgggag	1320
agcctgcagc	tgcctgtcag	ccccagctcc	acaaacagct	acaggtgttc	ttcccgagat	1380
ggagagtcca	ctctccaccac	actgaaaaag	gaggccaaat	ttgagctgag	tgccctccat	1440
gaggacaaga	agccctccaa	gctctttgag	gatgacgagc	atgagaaaga	caactactgc	1500
attgaaaaag	tgagcgcttc	agaggagatg	ctggagctgg	aaaaggagag	gagagagctc	1560
atcccgacgc	aggccgtcaa	gaagaatcct	ggcattgcag	caaaatggtg	gaatccccc	1620
caggaaaaaa	ccatcgaggga	gcagctggac	gaggaaatc	tggagtgcga	caaaaagta	1680
aaggagcgca	aagagagaag	ggcacagcag	gaacagttgc	tgctgcagaa	gcagttacag	1740
cagcagcagc	agcagccccc	atcgcagctc	tgcacagccc	ctgctctctc	tcagtaagcc	1800
gcagcatgca	ttgacaaaag	aaaggaggac	attgtcacag	agcagataga	ttctctctgt	1860
gctcgcaaac	aatttcagct	gatggagaat	tcacggcgaag	cggtggccaa	ggcccgagat	1920
acaccccaggc	tgtcttccat	caagcctttc	tacaggcctc	tggggtcagt	caactcagac	1980
aagccactga	ctaactcogag	accaccttct	gtcgggggac	ctccagaaga	cagtggtgac	2040
tcagccgccca	agggacagaa	atcccccggt	gcctcggaag	ccccctcggc	agcagggaag	2100
caggccaaca	cagcctctca	ggggaaggaa	ggccctacac	gcgagcctc	taaacgtggg	2160
cccttatcta	aactgtgggc	tgaggatgga	gaatttacga	gcgcccgggc	tgctctcaat	2220
gtggtcaagg	atgatgacca	tgggattttg	gatcagttct	caagatctgt	taactgtctc	2280
ttgacccaag	aggagcttga	ctctggtctg	gacgaattgt	cggtgagctc	tcaggatacc	2340
acagtccagg	agaccctatc	caatgatctc	agcatgcaca	acatcagtag	cagcggggca	2400
tcacatgaga	caaccaatgc	ctccaggaaa	aattcaactg	ctgatttttc	ttcgtcccag	2460
acccacaaga	ctgacaaacc	ctcagagggc	cgaggaaag	gcgtctccaa	gattcttagt	2520
gatcatgtgt	tctattcccc	ttcctccacg	ctgggggaat	ctcogttggt	tgatgacccc	2580
ttggagtatc	aggctggcct	ctggtgcag	aatgccattc	aaacagccat	agccggagcag	2640
gtggataaag	ctgtgtccaa	aaccagcagg	gatggagcag	agcaacaggg	acctgaagcg	2700
actgtagagg	aagctgaagc	tgccgttttc	ggctcagaaa	agcctcagag	catgttttag	2760
ccactccagg	tgtctctccc	tgttcaagag	aaaagggaag	tattaccaaa	gattctgcct	2820
gctgaagaca	gggcgctcag	ggaagggggg	ccccccacgc	cactgcaccg	tgtgcagccc	2880
agtggcccgga	ttaaactgga	ggagaccagg	cccgaaggaa	gctatttcag	caagtactcg	2940
gaggcagctg	agctgagaag	cacagcctcc	ctcctggcca	ctcaagaatc	tgacgtgatg	3000
gttgggccttt	tcaagcttag	gtccaggaaa	cagcggactt	tgtccatgat	tgagggaagc	3060
atcogagcag	ctcaggaaa	ggaaggaggag	ctgaaggagg	agagacaagt	cttgcagagt	3120
acgcagagccc	ccaggacaaa	gaatgcccca	tcactgcctc	ccagaaactg	ctacaaaact	3180
ctgcacaggga	aaatgagaaa	agtcaaacct	ctccatcccc	ccaccactga	agggcccgag	3240
ctcagcgtct	acttagcccc	tgaagaggct	gccggaaacc	agcggcccaa	gaatctgatg	3300
cagacccotca	tggaaagacta	tgagacacac	aaatctaaaa	ggcgcgagag	aatgatgatg	3360
agtgtgtaca	cttctaagtt	actgtcttgc	aaagtgactt	ccgaggtctt	cgaggccaca	3420
cggtgttaac	gaagaaaagag	cgactggctt	ttgcgtcggg	agcagggagt	ctatgccaac	3480
caggagggaag	aagacaaacga	ataaacttcc	ctcaaccagc	gaagcgtctt	tggtgtcttg	3540
gagacccaaga	aaaccaagaaa	ttaaacaactg	aaagcatttt	aatggactat	ttattaaagt	3600
gcaaacaaaac	taagcaatcc	tattgtagac	cagaagctgc	aatatacaat	gatgaaaagt	3660
aaacgaaaag	gaagtccccc	ctatgacact	tggaacccca	ggaatcaaaa	gagaaaggac	3720
tttttttggt	actccaatccc	agtgtctggt	ttggggccaa	tctacaattg	atccaaaagg	3780
acaaaacatg	tctcaagcaa	atcaacatag	ttgtctcagg	cggaactcag	accagcctct	3840
ctgggcgatg	aaggacaatg	gtggttagggc	cacttagaga	gcagtagggc	ccaggtctgt	3900
gcagcctaca	tgttccagga	tatttccctg	caccagtttc	ccacctgtca	cttgcagcat	3960
ctgggggtgt	tgcattaaat	ccattttctc	cagcagccat	aaaatgaatg	gcgtttctct	4020
ccagatttgc	ccaggaaggg	cagttatgat	tccacttctc	ctacatttaa	ttatttttaa	4080
ttgagacaaca	ctcatgatca	caattcaata	aatccattat	ttagacaat	agttctctat	4140
gtcaaaaaca	ccatctattc	tacatgagga	ggctgagaaa	ggggaagaga	aaaaacaaac	4200
aaatggatga	tgttgccttt	aggatcgtgt	ctgcttataa	tattcaactga	ttcgcagttt	4260
atgagacac	attttccacg	ttatgtttct	tcataggaaa	actatattta	gtgggcaaca	4320
actattttta	tgatgggatg	ggggagtata	tacaogtata	gaatctgtac	gcgttgaaac	4380
acttgggtca	agatgggtgg	ggcattttta	gagcggcaat	aattgaaaaa	aaaggcgaa	4440

ttctgcttgg	agaggtagat	gataaagaat	aaaaaggtgt	ttataactat	tttgtattat	4500
aaagtggggc	ttagagatag	gaagaagaat	gatggattcc	ttttggatca	atcagaaagg	4560
aaacacgaaa	gaaaagtcag	gaaggtagag	agagaaaaag	ggaggggaag	agaaagaagt	4620
ggaataaaat	aaggaggtaa	gagatactat	ttttgctgag	caaccagttg	gtttcaggat	4680
gatacaaaag	aaaatataga	atagaaataa	gtgcaggctt	ggaatcagct	acaaatccta	4740
aaagtggggg	gtgtgtggat	gtgtgtgtgt	gtgtgtgtac	accattgtgt	gtttgtaaaa	4800
ttgttatgtt	catgagtaag	ggtgtgtgtg	ttgtgtgtatt	aaaattccag	agtgaccgtg	4860
gcacttgggt	gtacaggttaa	ttcctccaga	gctgttttgt	ggcttcaggga	gtggagtgag	4920
aattttcttt	ttatgaaaag	ggatataaag	gcaccgagct	gatgcagtat	ttgtaatatt	4980
aagttgacct	acaacggtat	ttgcatgagt	cacaattaca	aagttttgag	cggtttttga	5040
atttgacatt	taggaaagtc	tcctatttat	ttctatactt	tacattcatg	cttagtatac	5100
tatagaggat	gccagcttta	atctttctgt	catttaaagc	aatatgataa	gggtattcaa	5160
taattgggtg	ccctaaattt	ctggatgaga	aaattttcaa	ttctggccag		5210

<210> 419
 <211> 449
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig63102_RC

<400> 419						
ctatcccaatt	ttttcattta	tttttatgag	gagcctcaaa	aaatagaagt	ggccttggtg	60
gggaaggaat	gtcgtgattc	tgagtacagt	atggcaaat	cttttgccct	agttcccta	120
actccatcca	gtgcatatga	aaatgtctat	gtgcgtacaa	ggtaggtctg	gactaagctc	180
tgccctcccg	gtgtgtttcc	cattgcggga	gcttaagcca	gggtgcccg	gactcctgtg	240
tggtctgttg	gcctgatgcc	cacagctgag	gattcaggaa	ggaggggctg	gagctctgct	300
gcagactctc	cgcgccctca	cagcccaaac	cgggcaactg	agcaatgagg	tggtcagctc	360
agggagagggt	cgcgaagtt	ctcgccacaa	tgctgcaggga	agtcaaagtc	agggcaggag	420
aaaggcagcg	gggacacatt	tttggcctg				449

<210> 420
 <211> 3044
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003981

<400> 420						
gtttgcgggt	ggttgtgtgt	ctcggggcgg	tggtggagtag	gtctggacct	ggactcacgg	60
ctgcttgagg	cgtccgccat	gaggagaagt	gaggtgctgg	cggaggagtc	catagtatgt	120
ctgcagaaag	cctcaaatca	ccttcgggaa	atatgggagc	taattgggat	tcacagaggc	180
cagcgggtac	aaagaactga	ggttgtaaag	aagcatatca	aggaactcct	ggatatgatg	240
attgctgaag	aggaagacct	gaaggaaaag	ctcatcaaaa	gcataatccg	ctgtcagaaa	300
gatgtgaaca	ctctgtgcag	cgagttacat	gttgagccat	ttcaggaaga	aggagagacg	360
accactctgc	aactagaaaa	agatttgccg	accocagggt	aattgatgag	aaaacagaaa	420
aaggagagaa	aacaggaaact	gaagctactt	caagagcaag	atcaagaaat	gtgcgaaatt	480
cttttgatgc	ccactatga	tattgacagt	gocctcagtc	ccagcttaga	agagctgaac	540
cagttcagtc	aacatgtgac	aactttgagg	gaacacaaag	cttctaggcg	tgagtgagtt	600
ctcagtataa	agagacagat	catactgtgt	atggaagaat	tagaccacac	ccagacacac	660
agctttgaag	gacagtgtgt	gtgtgaagac	gaagatgcct	tttgtttgtc	tttgagagat	720
attgcaacac	tcaaaaagtc	gctacggcag	ctggaaatgc	agaaatcaca	aaatgaagca	780
gtgtgtgagg	ggctgcgtac	tcaaatccga	gagctctggg	acaggttgca	aatacctgaa	840
gaagaaagag	aagctgtgac	caccattatg	tctgggtcaa	agggcaaggt	cgggaaagcg	900
ctgcacattag	aagtggatcg	gttggaagaa	ctgaaaaatg	aaaacatgaa	gaaagtgtat	960
gaggcaattc	gagtgagact	ggttcagtag	tgggaccagt	gcttttatag	ccaggagcag	1020
agacaagctt	ttgccctttt	ctgtgctgag	gactacacag	aaagtctgct	ccagctccac	1080

gatgctgaga	ttgtgcggtt	aaaaaactac	tatgaagttc	acaaggaact	ctttgaaggt	1140
gtccagaagt	gggaagaaac	ctggaggctt	ttcttagagt	ttgagagaaa	agcttcagat	1200
ccaattcgat	ttacaaaccg	aggaggaat	cttctaaaag	aagaataaac	acgagccaag	1260
ctccagaaaa	tgtctgccaa	gctggaagaa	gagttgaagg	cacgaattga	attgtgggaa	1320
caggaaacatt	caaaaggcatt	tatggtgaat	gggcagaaat	tcattggagta	tgtggcagaa	1380
caatgggagga	tgcattcgatt	ggagaaaagag	agagccaaagc	aggaagagca	actgaagaac	1440
aaaaaacaga	cagagacaga	gatgctgtat	ggcagcgctc	ctcgaacacc	tagcaagcgg	1500
cgaagactgg	ctcccaatc	acggggcaaa	gcacgtaagc	tgaacactac	caccatgtcc	1560
aatgctacgg	ccaattagtag	cattcggcct	atctttggag	ggacagtcta	ccaactcccc	1620
gtgtctcgac	ctctctcttc	tggcagcaag	ccagtcgctg	cttccacctg	ttcagggaag	1680
aaaacacccc	gtactggcag	gcattggagcc	aacaaggaga	acctggagct	caaaggcagc	1740
atctcgatgtg	gtgggtaccc	tggctcggcc	ccccccagc	gcaacttcag	cattaattct	1800
gttgcagcga	ctattctga	gtttgcgaag	gatccgtccc	tctctgacag	ttccactgtt	1860
gggcttcagc	gagaagcttc	aaaggtctcc	aaatctgatg	ctactctcgg	aaactccaat	1920
tcaacccaaca	tccagtcctg	agaagccctg	atcagtcac	cagctgtggc	ttcctgtgcc	1980
tagactggac	ctaattatat	gggggtgact	ttagtttttc	ttcagcttag	gagtgcttga	2040
aaacttggcc	aggttccatg	accatgggccc	taacttaaa	atgtgaatga	gtgttacagt	2100
tgaagcccca	tcattagttt	agtgtgccta	ggagacttgg	ttttgactta	tatacatgaa	2160
aagtttatgg	caagaagtgc	aaatttttagc	atatggggcc	tgacttctct	accacataat	2220
tctacttgct	gaagcatgat	caaagcttgt	ttttttccac	cactgtagga	aaatgatgtg	2280
ctatgcccat	ccctgggggt	aattttggca	tgtatacctg	taactagtaa	ttaacatctt	2340
ttttgttttg	gcattgttcaa	ttaattgctgt	agctatcata	gctttgctct	tacctgaagc	2400
cttgcoccca	ccacacagga	cagccttctc	cctgaagaga	atgtctttgt	gtgtccgaag	2460
ttgagatggc	ctgccttact	ggcaaaagag	tgacaggaag	gctgggagca	gctttgttaa	2520
attgtgttca	gttctgttac	acagtgcatt	gccctttgtt	gggggtatgc	atgtatgaac	2580
acacatgctt	gtcggaaagc	tttctcggcg	ttgtccctct	ggctctcact	tccccattc	2640
ctgtgcctacc	tttgcctgag	ttctcttacc	ccgcagttg	ccagccagat	tgggagctgt	2700
tttgttccaa	tgggttgagc	tgtctttgtc	gtggagatct	ggaactttgc	acatgtcaact	2760
actggggagc	tgttctcgtc	ctagcttcca	cgatgagcgc	ccctctttac	ctactctctc	2820
aatcactact	ctttctgaag	cactattatt	tattctctcc	ctgtctgcct	gcagcagtag	2880
tactgtcaac	atagtgttaa	tggttctcaa	aagcttaacca	gtgtggactt	gggtgttagc	2940
acgctgttta	ctctatcacg	tacgtgtcct	gttttttaaaa	tatacaatta	ttctttaaaa	3000
taaattaaaa	tctgtatact	acatttcaaa	aaaaaaaaaa	aaaa		3044

<210> 421

<211> 407

<212> DNA

<213> Homo sapiens

<300>

<308> Contig20217

<400> 421

agtgttacaag	acagataattg	tgaatttttt	ccttgggctt	ctttcaattt	ggttcaattc	60
tgttccagta	ttctcataac	atttggatca	ggtaacttat	caactgttag	tgtcactgtt	120
ccctttctgta	tctctcggaa	gcctcctgtt	actcgttttc	taattctctt	ccagcccaat	180
gcaagatcta	gtgtcgaaag	gatgaacacg	ctgagcagaa	cagttcttta	gtgttcggcc	240
taaatgacca	acttcattgac	ctctcaaggt	ctcgtccccc	ggctgcacgc	agctcagctt	300
ccaacattcg	ggaaaggcgt	gcgggatcgc	gaggggccgc	tgggaccgag	tccggacgtc	360
gcgacgcgca	gggggctgct	gggcccgcgc	tctcgggaaa	agtggag		407

<210> 422

<211> 1170

<212> DNA

<213> Homo sapiens

<300>

<308> Contig2399_RC

<400> 422

aatgtttcaa caacatttat acatcttttt aattagatta gctttacaag caacttgaga 60

```

gctctttcat aatgaaacac tgcctttcaa attacatgac tttatagtct atctgtatat 120
gaactcctgaa aatctttggcc tactcttgat aaaggtctgt gttttacage tgggttaagt 180
gttgaaataaa aacagcatta tgaagcttta tctttaacca ggattaaacc catataggcc 240
acaaacagtt ttaaattttg tttgaaatag agtctcgtct ggtgacaaagt taatagtccc 300
actaacatca aatgaatgga tattaataaa atggccatcc tgattttctg atcttttacc 360
aatgatctct aagagagtat ataaactttt ccaaaaaata tcaattttaa gctcagctctg 420
gcaagacaaa cccacagata gtagaatcta atacttttta caactgtaga ggtacacatt 480
tgaatgacaa cagggtctcc caaattaaaa ggctgtggct gtacattttg gtacataaac 540
aaactcccca tcaacataaa acaattttcta ctaaccctaa aagaataaat gcttttgagt 600
tgtggctcaa gcggcgccac tgtactatag accagatggc taacacactt gcgctctcgt 660
ggttgggctg gcagaggcac ttaactgtaga gatgtgatgt tatcatctgg aaggaaaggg 720
aggaagttaga acaaaactga atatggcgca ggaaaaaaca ataaactatc atcttctcct 780
gcaaattctc ctaagtgtaga tcaactatct gaacattcaa tcccatccta agtccaactgt 840
tggctctcaa acacaatttt ttttccatgt ttacatgaa tagcaatggc aatcttttaa 900
taaaaattac ttagcttccc atagaaaagg tctcaaatgt aatacaaaact atacagatgt 960
taaaatgtct ttaagttgaa atatacagaa aattttctga tactcaact ttggcaaga 1020
agatcaagga cagagataga aaacctaaat gggcaaggtc ttacaccat aattcacgca 1080
aaatatggat acaacccttt aggttttgcg aaaaagccta ggaaagtga gctcaagaa 1140
tcgctcccaa ttcattcttt cctcgtgcg

```

<210> 423
 <211> 2019
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007036

```

<400> 423
tgggtcacgg ctgcttccca ccagcaaaaa ccacgactgg agagccgagc cggaggcagc 60
tgggaaacat caagagcgtc ttgctgctga ccacgctcct cgtgctcgca cctggtgtgg 120
ccgctctggag caataattat gcggtggact gccctcaaca ctgtgacagc agtgagtcca 180
aaagcagccc gcgctgcaag aggcacgtgc tcgacgactg tggctgctgc cgagtgtgag 240
ctgcaggggcg gggagaaact tgcctaccga cagctctcagg catggatggc atgaaagtgtg 300
gcccggggct gaggtgtcag ccttctaagt gggaggaatc ttttggtgaa gagttttgta 360
tctgcaaaaa ctgtccctac ggcaccttcg ggatggatgg cagagagacc tgcaactgcc 420
agtcaggcat ctgtgacagg gggacgggaa aatgcctgaa atccccctc ttccaattat 480
cagtaaccaa gtcttccaac agatttggtt ctctcacgga gcatgacatg gcatctggag 540
atggcaatat tgtgagagaa gaagtgtgta agagaaatgc tgccgggtct cccgtaatga 600
ggaatatggtt aaatccacgc tgatcccgcc tgtgatttct gagagaagcg tctatttttcg 660
tgattgttca acacacagcc aacattttag gaactttcta gatatagcac taagtacatgt 720
aatttttgaa gtccaaaatt gtgatgcagt gtggatccag aaaacaaaaa gtgagatact 780
tacaatccat aacatccata tgactgaaca ctgtatgtg tttgttaaat attcgaatgc 840
atgtagattt gttaaattgt tgtgtatagt aacactgaag aactaaaaat gcaattttag 900
taatcttaaa tggagacagg tcaacaaaag agggagctag gcaaaagctg agaaccaggt 960
gaagcaaaat agtcttttga ctttgatgta ctttaattgt gggatagga atgaagact 1020
aagagcagga gaagatgggg aggggggtgg agtgggaaat aaaaatttta gcccttctct 1080
gttaggtagc ttctctagaa tttaatgtgt cttttttttt ttttttggct ttgggaaaaa 1140
tcaaaataaa acaaccagaa aacctcgtga ggaagtgaaga ttttgaagc ttatggaaat 1200
ttgagttaaca aacagcttgg aactgagagc aatttcaaaa ggtgctgat gtagtcccc 1260
gtttacctgt actcgaagga cggttctcgg gcatagggaa acataacact tccataaata 1320
gcttttaactg atgccacctc agagataaat ctgaagaata ttttaccac ttggtggttg 1380
tgtgtgtagt aagtgaaata tttatatatt ttataaata atgtgtgtg tgcaagctat 1440
cttccctacc catatttatc atcctcttga gaaagaataa ctagtattat ttgttgaaa 1500
tgttttagat aaaaacctat gactctataa ggttttcaaa cactgtgagc atgtgaaatt 1560
tattatccat aattatagga gtcaactctg atttcaaaaa atgtcaaaaa atgagcaaca 1620
gggagcactt atttaaacat aagtgtctgt acttcggtga attttcaatt taaggtatga 1680
aaataaagtt ttgaggagtt tgtaaaagaa gaatacaatt tcacagaaaa acatgtcaac 1740
tttaaaatct aggttggaatt aggtatatat ttgaaaagat cttagcaca acagagctgt 1800
tgtactagat gttcttagga aatatctcag aagtatttta tttgaagtga agacttatt 1860
taagaattat ttcagtattt acctgtattt tattcttgaa gttggccaac agagtgtgta 1920

```


atgtgtgttg aagcccttg aatgtaaac tgcataagct gttagggttt gttttaaaag 1980
gacatgttta ttattgttca ataaaaaaga acaagatac 2019

<210> 424

<211> 1720

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018354

<400> 424

gactgcagat gaaattagta actggtgggg tcgtgggggt tgaatgggtg gcggggagcag 60
ctatgtcagt tgggtgtgttt ctgcttatgt tagggtaatt gggcacggcc ttgtgttaac 120
tgggtgaatat ctctgaacct gggcatgaaa cagagagatg tcttaactct ggggtgagagg 180
aatctcatt tttctctgcc ctctcactgt ggcatactaa gaaaaaagt ttgggttccct 240
cgagcatgaa ggagagctct gctcccaaaa tttgggagct ccagatttct tccagggtgt 300
ggaggcatca atatatcagt ctgggaaaag ggttctctgg ccactccagg agctgagttg 360
gggtggaaggt gctgagagtg tgggtggggg ccacttctga gcaacctgt ggcaccacct 420
gctgtgtccct gttgtggct gggcaactcag gaaaatgttt ttggtgctaa gagtaaaaag 480
ccaaccaaca aacacatctc ttttttctgt ctattcactg gaaagtaaaa gcagctctggg 540
cgcaggtctg ggaccagat ggaattcaaa cttatgctgt ctctcaaggt gctcacggtt 600
gctgataaac agctggataa aatgaagagt ctatgagtga gggatgcaga gccagggaag 660
gctggtggag tgatgcacac agcacagggg tatgagttg cagctgccaa ggggccaaag 720
gatgagctgg ggccctcctt cccaatggca tctcccctgt gtcctggaact gaagacactg 780
agcaatggtc cccaagcccc aaggagatca gctccccctg gcccagtgcc cccaaccagg 840
gggtgtgtgg agaattgcctg ctctctctca gaggagcatg agaccattt ccagaacct 900
gggaacacga gactggcgag ctccaccagt ccccctgggg gtgtctctc actgcccaga 960
tcccagcggg atgatctgtc ccttcatcca gaggaggggc cagccctgga gccctgagc 1020
gcgccgtgtg attatggctt tgtttccgct ctgttttctc tgggtgagtg gattctctg 1080
gtggtgcagc catagcccat ccccctgtag gctcgagtca atccggacac agtgacagcg 1140
cgggagatgg aacgactgga gatgtactac gccgcctag gctccactgc agtgacagtg 1200
atcatcgag gctctgggct gctcacgggt gctcacggc ggcggcatgc tcttctcggt gctgctcatg 1260
gtctccctgt gcaaggcgga gctgtaccgc tgcgtacccg cagaggacct tgcgtcccgg caagggtctc 1320
aggaagacct acggtcccat taacctgcgc atgagacagc tcaatgggga tggggggcag 1380
gcctcggttg agaatgaagt tgtccaggct tcagagacta gccacacct ccagggtct 1440
taagaactag cccaccttat ctggctgctt tagctccagt gctacaaggt ccaccacctg 1500
tcccggccca cctgacctt gcccaaggccc tggggtttta aactgagctc acatagggcc 1560
ttgtggaaga agtactgggt gctggaggga gagctcgggg cccagcccat gccccacacg 1620
ggcaagcagc ccaactgact gttttgtagc tgagggtttg catacggttt tgtttggagg 1680
atggctcttg ctgtcaaaaa tacaaaaagt tggaaacgcg 1720

<210> 425

<211> 987

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000125

<400> 425

cgagaggggg gagaacagac aacgggcggc ggggagcagc atggagccgg cgccggggag 60
cagcatggag ccttcggctg actggctggc cagcccgccg gcccggggtc gggtaggga 120
ggtcgccggc gctctggagg cgggggcgct gcccaacgca ccgaatagtt acggtcgagg 180
gcgcatacag gctatgatga tgggcagcgc ccgagtggcg gactgctgc tgctccacgg 240
cgccggagccc aactgcgccg accccgccac tctcaccgca cccgtgcagc acgctgcccg 300
ggaagggcttc ctgggacacg tgggtgtgct gcaaccggcc gggggcgccg tggactgtgc 360
cgatgcctgc ggccgtctgc ccgtggaact ggcgtgaggag ctgggcccac gcgatgtcgc 420
acgtacctgc cggcgcgctg cggggggcac cagagggcag aacctgccgc gcatagatgc 480
cgccgaaggt cctccagaca tcccggattg aaagaaccag agaggtctcg agaaacctcg 540
ggaaccttag atcatcagtc accgaaggtc ctacagggcc acaactgcc cgcgccacac 600

```
ccaccccgct ttogtagttt tcatttagaa aatagagctt ttaaaaatgt cctgcctttt 660
aacgtagata taagccttcc ccactaccg taaatgtcca tttatatcat tttttatata 720
ttcttataaa aatgtaaaaa agaaaaacac cgcttctgcc ttttactgt gttggagttt 780
tctggagtga gcactcaccg cctaagcgca cattcatgtg ggcatttctt gcgagcctcg 840
cagcctccgg aagctgtcga ctcatgaca agcattttgt gaaotagggg agctcagggg 900
ggttaactggc ttctcttgag tcacactgct agcaaatggc agaaccaag ctcaataaaa 960
aataaaaataa ttttcattca ttactc 987
```